Dynamic Pricing and Inventory Management in the Presence of Online Reviews

Nan Yang
Miami Herbert Business School, University of Miami, nyang@bus.miami.edu

Renyu Zhang
CUHK Business School, The Chinese University of Hong Kong, philipzhang@cuhk.edu.hk

We study the joint pricing and inventory management problem in the presence of online customer reviews. Customers who purchase the product may post reviews that would influence future customers' purchasing behaviors. We develop a stochastic joint pricing and inventory management model to characterize the optimal policy in the presence of online reviews. We show that a rating-dependent base-stock/list-price policy is optimal. Interestingly, we can reduce the dynamic program that characterizes the optimal policy to one with a single-dimensional state-space (the aggregate net rating). The presence of online reviews gives rise to the trade-off between generating current profits and inducing future demands, thus having several important implications upon the firm's operations decisions. First, online reviews drive the firm to deliver a better service and attract more customers to post a review. Hence, the safety-stock and base-stock levels are higher in the presence of online reviews. Second, the evolution of the aggregate net rating process follows a mean-reverting pattern: When the current rating is low (resp. high), it has an increasing (resp. decreasing) trend in expectation. Third, although myopic profit optimization leads to significant optimality losses in the presence of online reviews, balancing current profits and near-future demands suffices to exploit the network effect induced by online reviews. We propose a dynamic look-ahead heuristic policy that well leverages this idea and achieves small optimality gaps which decay exponentially in the length of the look-ahead time-window. Finally, we develop a general paid-review strategy, which provides monetary incentives for customers to leave reviews. This strategy facilitates the retailer to (partially) separate generating current profits and inducing future demands via the network effect of online reviews.

History: Submitted in July 2019; revision submitted in December 2021; accepted in February 2022 by Qi Annabelle Feng.

Key words: dynamic pricing and inventory management; online customer reviews, network effect

---

1. Introduction

[1]Most e-commerce marketplaces (e.g., Amazon, Taobao, eBay, etc.) have adopted the crowd-sourced online review systems that facilitate customers to rate products purchased through the platform. In

[1] The paper was previously entitled "Dynamic Pricing and Inventory Management under Network Externalities".

these marketplaces, customers see such a wide variety of products that they are typically uncertain about the quality and fit of each individual product. The customer-generated reviews/ratings provided by online platforms reduce such quality uncertainty concerns of potential customers. As a consequence, the online reviews/ratings enable customers to make better-informed purchasing decisions, thus increasing customer satisfaction and decreasing post-purchase regrets. From the perspective of the merchants on the platform and the platform itself, the reviews and ratings provided by previous customers could help them more accurately forecast future demand (Dellarocas et al. 2007), and, thus, improve their operations strategies such as pricing and inventory control.

Potential customers' response to online reviews/ratings generated by previous buyers has two important features. First, customers are more likely to purchase the product that has more reviews and higher ratings. On one hand, more reviews convey more information about the product, thus more effectively reducing information asymmetry and removing customers' concern about quality uncertainty (see, e.g., Papanastasiou and Savva 2017). On the other hand, more reviews and higher ratings demonstrate higher popularity of the product, and hence signal a better quality as well. To increase its traffic, the platform would also recommend the product with a larger volume of reviews and a higher rating to new customers by, e.g., highlighting these products on the homepage of the platform website. Second, customers take the recent reviews more seriously than reviews generated long ago[1]. It is natural that a customer would consider more recent reviews more relevant. Furthermore, most platforms display reviews and ratings in an (almost) reverse-chronological order, with more recent reviews on higher positions. As we will show later in this paper, both features of online customer reviews/ratings have profound implications on the operations strategies of a merchant on an e-commerce platform and the platform itself.

It is clear from our discussions above that customer-generated online reviews and ratings would substantially impact the purchasing decisions of potential customers, thus influencing the demand and sales of the product. Furthermore, the sales of the product would in turn impact reviews and ratings for the product. Such inter-correlations between demand/sales and user-generated online reviews undoubtedly pose significant challenges for merchants and platforms to optimize their operations (pricing and inventory) strategies. On one hand, the pricing and inventory decisions should respond to the online reviews to leverage their significant impact on the demand of a product. On the other hand, the price and available inventory of a product would influence its sales, thus reshaping the online reviews and ratings. Furthermore, it is also a common strategy for online retailers to pay customers to leave reviews, which could help expand their future demand[2].

[1] See, e.g., BrightLocal 2017, `https://www.brightlocal.com/learn/local-consumer-review-survey/`

[2] See, e.g., Green 2017, `https://www.kayako.com/blog/asking-for-reviews/`

Motivated by the aforementioned challenge, we study the dynamic pricing and inventory strategy in the presence of online customer reviews. We analyze a single-item periodic-review joint pricing and inventory management model, incorporated with customer-generated product reviews/ratings. More specifically, in each period, a fraction of the customers who make a purchase would post a review. The review may be either positive or negative. The probability that a customer will leave a positive (resp. negative) review is higher (resp. lower) if the product she requests is available, since she receives a better service in this case. To model the overall impact of buyer-generated reviews, we define the *aggregate net rating* of the seller/product as the weighted sum of the differences between the number of positive reviews and that of negative ones throughout the planning horizon. Furthermore, a new customer's willingness-to-pay is increasing in the aggregate net rating of the seller/product. The online aggregate net rating also captures the effect that recent reviews have stronger impact on customers' purchasing decisions than reviews generated long ago. In this model, we characterize the optimal pricing and inventory policy of a profit-maximizing e-commerce firm in the presence of online reviews. Our analysis highlights the significant impact of buyer-generated reviews/ratings on the firm's optimal price and inventory policy, and identifies effective strategies to leverage the demand-shaping effect of online reviews. We also remark that our model can be easily adapted to study the joint pricing and inventory management of other products that exhibit strong network effects (e.g., game consoles). For such products, customers would be attracted by other customers who adopt the same product so that the firm needs to balance current profits and future demands carefully, as with customer-generated reviews on an e-commerce platform.

1.1. Main Contributions

To the best of our knowledge, we are the first in the literature to operationalize online buyer-generated reviews/ratings in an inventory model and study their impact upon the joint pricing and inventory policy of a firm. We show that a rating-dependent base-stock/list-price policy is optimal. Moreover, we make an interesting technical contribution: The dynamic program to characterize the optimal policy can be reduced to a single-dimensional one (aggregate net rating). We perform a sample path analysis and show that, under the optimal policy and low initial inventory level, the inventory dynamics of the firm have no impact on the optimal policy with probability one. Such dimensionality reduction significantly simplifies the analysis and computation of the optimal policy, and thus helps deliver sharper insights on the managerial implications of online reviews/ratings.

Our analysis reveals that online product reviews lead to a network effect: Customers are more willing to purchase the product with higher previous sales. Such network effect drives the firm to balance the trade-off between generating current profits and inducing future demands and, thus, has a few important and interesting managerial implications. First, online buyer-generated reviews give rise to the service effect and rating-dependent pricing: The firm provides better services to customers in the presence of online reviews, and charges different prices with different ratings. Customers who are satisfied immediately (i.e., not being put onto the wait-list) are more likely to post reviews. Therefore, if benchmarked against the case without online reviews, the firm provides a better service to customers. Hence, the safety-stock and base-stock levels are both higher in the presence of online reviews. Online reviews also have significant impact upon the firm's pricing policy. If the aggregate net rating is low, the presence of online reviews decreases the sales price to induce higher future demands. Otherwise, if the aggregate net rating is high, the firm increases the sales price to exploit the higher potential demand. From an intertemporal perspective, the firm should put more weight on inducing future demands at the early stage of a sales season. Thus, when the market is stationary, the firm charges lower prices at the beginning of the planning horizon. Hence, the widely-adopted introductory price strategy (offering price discounts at the beginning of the sales season of a product) may stem from the presence of the online review system.

Second, the aforementioned trade-off between generating current profits and inducing future demands drives the stochastic rating process to follow an interesting mean-reverting pattern. As long as the firm adopts the optimal joint pricing and inventory policy, the aggregate net rating will increase (resp. decrease) in expectation when it is below (resp. above) the "mean" currently. When the rating is low, the firm underscores inducing future demands, so the aggregate net rating will have an increasing trend. On the other hand, with a high rating, the firm simply extracts high profits from the high potential demand, so the rating will fall in expectation.

Third, although myopic profit optimization leads to significant losses in the presence of online reviews, balancing current profits and near-future demands suffices to exploit the effect of buyer-generated reviews. Our extensive numerical studies demonstrate that the myopic policy ignores the (inter-temporal) demand-inducing effect of online reviews, so it substantially erodes the firm's profits. We propose a look-head heuristic policy that dynamically maximizes the profit in a (short) moving time-window. Interestingly, this heuristic policy yields small optimality gaps (less than 2% in our extensive numerical experiments) with exponential decay in the length of the moving time-window. Thus, with the dynamic look-ahead heuristic, the firm can effectively leverage customer-generated online reviews by balancing current profits and *near-future* demands.

Finally, we demonstrate the value of paying customers to review the product. It is well documented that online retailers sometimes pay customers to review their products[1]. The key idea underlying such a paid-review strategy is that, the firm employs an additional leverage (i.e., incentives for customers to leave reviews) to partially separate generating current profits and inducing future demands through

[1] See, e.g., Green 2017, `https://www.kayako.com/blog/asking-for-reviews/`

net ratings. With sufficiently strong impact of net ratings on customer demand, it is optimal for the firm to pay customers for reviewing its products, regardless of its inventory level. The optimal price is higher, whereas the safety-stock and base-stock levels are lower if the firm adopts the paid-review strategy. In other words, the firm which pays customers to leave product reviews is able to charge a premium price for the product and maintain a low inventory level to generate higher current profits.

The rest of this paper is organized as follows. In Section 2, we position this paper in the related literature. Section 3 presents the basic formulation, notations and assumptions of our model. Section 4 simplifies the model to a single-dimensional dynamic program. We investigate the key trade-off in the presence of online reviews in Section 5. We discuss the extensions with paid reviews and with high initial inventory level in Section 6. Section 7 concludes this paper by summarizing our main findings. All proofs are relegated to the Appendix. We will use $\mathbb{E}[\cdot]$ to denote the expectation operation. In addition, for any $x, y \in \mathbb{R}$, $x \wedge y := \min\{x, y\}$, $x^+ := \max\{x, 0\}$, and $x^- := \max\{-x, 0\}$.

2. Literature Review

The literature on the joint pricing and inventory management problem under stochastic demand is rich. Petruzzi and Dada (1999) give a comprehensive review on the single period joint pricing and inventory control problem, and extend the results in the newsvendor problem with pricing. Federgruen and Heching (1999) show that a list-price/order-up-to policy is optimal for a general periodic-review joint pricing and inventory management model. When the demand distribution is unknown, Petruzzi and Dada (2002) address the joint pricing and inventory management problem under demand learning. Chen and Simchi-Levi (2004a, b, 2006) analyze the joint pricing and inventory control problem with fixed ordering cost. They show that (s, S, p) policy is optimal for finite horizon, infinite horizon and continuous review models. Chen et al. (2006) and Huh and Janakiraman (2008), among others, study the joint pricing and inventory control problem under lost sales. In the case of a single unreliable supplier, Li and Zheng (2006) and Feng (2010) show that supply uncertainty drives the firm to charge a higher price in the settings with random yield and uncertain capacity, respectively. Gong et al. (2014) and Chao et al. (2016) characterize the joint dynamic pricing and dual-sourcing policy of an inventory system facing the random yield risk and the disruption risk, respectively. Yang et al. (2014) characterize the optimal policy of a joint pricing and inventory management model in which the ordering quantity must be of integral multiples of a given specific batch size. When the replenishment leadtime is positive, the joint pricing and inventory control problem under periodic review is extremely difficult. For this problem, Pang et al. (2012) partially characterize the structure of the optimal policy, whereas Bernstein et al. (2015) develop a simple heuristic that resolves the computational complexity. Chen et al. (2014) characterize the optimal joint pricing and inventory control policy with positive procurement leadtime and perishable inventory. Several papers in this stream of literature also integrate consumer behaviors. Yang and Zhang (2014) characterize the optimal price and inventory policies under the scarcity effect of inventory, i.e., demand negatively correlates with inventory. Chen et al. (2007) and Yang (2012) adopt the idea of constant absolute risk aversion and time-consistent coherent and Markov risk measure, respectively, to characterize the optimal joint price and inventory control policy under risk aversion. When the firm adopts supply diversification to complement its pricing strategy, Zhou and Chao (2014) characterize the optimal dynamic pricing/dual-sourcing strategy, whereas Xiao et al. (2015) demonstrate how a firm should coordinate its pricing and sourcing strategies to address procurement cost fluctuation. Yang and Zhang (2022) propose a new comparative statics approach to study a general joint pricing and inventory management model with market environment fluctuation and delayed differentiation. We have also adopted this approach to analyze the pricing and inventory model with online reviews in the current paper. Chen et al. (2016) develop a novel transformation technique to establish concavity for the joint pricing and inventory control problem in the presence of reference price effects. We refer interested readers to Chen and Simchi-Levi (2012) for a comprehensive survey on joint pricing and inventory control models. Building upon a standard joint pricing and inventory management model, we operationalize the buyer-generated online review system and study the impact of online reviews on the pricing and inventory strategies of a firm.

Online product review systems have received growing attention in the literature. For example, Dellarocas (2003) gives a comprehensive overview of reputation systems in general. Besbes and Scarsini (2016) show that, if buyers report the private-signal adjusted quality of the product, potential customers tend to overestimate the underlying quality of a product in the long run. Recently, several papers have studied pricing in the presence of the uncertain quality through different approaches to modeling the product review system. Crapis et al. (2017) study the monopoly pricing problem in which network effect arises from the fact that customers learn the quality of the product from their peers. Papanastasiou and Savva (2015) and Yu et al. (2015) study the impact of social learning (on product quality) when the product quality is uncertain and the customers are strategic. Shin and Zeevi (2017) propose an effective dynamic pricing policy in the presence of social learning and online product reviews. Adopting an evolution game framework, Mai et al. (2018) study how a ride-sharing platform could use bilateral ratings to control the behaviors of riders and drivers. Our paper differs from the above research by integrating online reviews into an inventory management model, and characterizing their impact on the joint pricing and inventory management strategy of a firm.

Another related stream of research is on the network effect (i.e., network externalities). In their seminal papers, Katz and Shapiro (1985, 1986) characterize the impact of network effect upon market competition, product compatibility, and technology adoption. Several papers also study dynamic pricing under network effect. For example, Dhebar and Oren (1986) characterize the optimal nonlinear pricing strategy for a network product with heterogeneous customers. Bensaid and Lesne (1996) consider the optimal dynamic monopoly pricing under network effect and show that the equilibrium prices increase as time passes. Cabral et al. (1999) show that, for a monopolist, the introductory price strategy is optimal under demand information incompletion or asymmetry. Recently, the operations management (OM) literature starts to take into account the impact of network effect upon a firm's operations strategy. For example, Hu et al. (2020) study whether a firm should reveal the sales information of a network product under demand uncertainty. Wang and Wang (2017) propose and analyze the consumer choice models that endogenize network effect. Under a newsvendor framework, Hu et al. (2016) propose efficient solutions for a firm to better cope with and benefit from the social influences between customers on online social media. Allon and Zhang (2017) characterize the optimal service strategies in the presence of social networks. Feng and Hu (2017) provide a unified theory that integrates the long tail and blockbuster phenomena in a competitive market under network effect. Shou and Zheng (2017) study the optimal pricing in a social network with strategic customers. In this paper, we characterize the network effect induced by buyer-generated online reviews/ratings, and show that such effect has significant impact on the pricing and inventory strategies of a firm.

Finally, from the modeling perspective, this paper is related to the literature on inventory systems with positive inter-temporal demand correlations (see, e.g., Johnson and Thompson 1975, Graves 1997, Aviv 2002). Our model is differentiated from those in this line of research with the following two salient features: First, we endogenize the pricing decision; second, the inventory decision (i.e., service level) influences the number of reviews posted by customers and, thus, the potential demand. Both features enable the firm to partially control the potential demand process. As a consequence, our focus is on the trade-off between generating current profits and inducing future demands in the presence of online reviews, whereas that literature focuses on the inventory control issues with inter-temporally correlated demands. The new perspective and focus of our paper facilitate us to deliver new insights on the managerial implications of online product reviews to the literature on inventory control with inter-temporal demand correlations.

3. Model Formulation

Consider a firm/seller who sells a product on an online e-commerce platform (e.g., Amazon or Taobao). We model the operations of this seller as a periodic-review stochastic joint pricing and inventory management system with full backlog. The focal sales horizon is T periods, labeled backwards as $\{T, T-1, \cdots, 1\}$. For most of our analysis, T is finite. We will also discuss the case with $T = +\infty$ when necessary. We assume that the platform adopts a crowdsourced online review system under which a customer can leave a review/rating for this firm or the product.

To simplify the analysis and clarify the model, we assume that each customer may either leave a positive review/rating or a negative one. The number of positive (resp. negative) reviews/ratings received in period t is denoted as r_t^+ (resp. r_t^-). This is consistent with the current business practice of e-commerce platforms such as Taobao. We use N_t to denote *aggregate net rating* of the firm at the beginning of period t, where

$$N_t = (\sum_{s=t+1}^{T} \eta^{s-t-1}(r_s^+ - r_s^-)) + \eta^{T-t} N_T \tag{1}$$

with the discount factor $\eta \in [0,1]$ and the effective number of reviews at the beginning of the sales horizon N_T. As shown in (1), our rating model captures two salient features of online review systems of e-commerce platforms. First, positive reviews would help the firm whereas negative reviews would hurt it. In our rating model, we use the difference between the numbers of positive and negative reviews to approximate the overall effect of all the reviews in the system. An alternative modeling approach is to use the proportion of positive reviews as the rating of the firm. We take the former approach because it is consistent with the rating system of Taobao[1], which is the largest e-commerce platform in China. Furthermore, our additive rating model also captures the effect that customers find the firm with more reviews more popular and reliable, and, thus, are more willing to make a purchase[2]. Second, customers view current reviews/ratings more relevant and informative than past ones. The discount factor η measures how customers value past reviews relative to current reviews. A larger (resp. smaller) η means that customers care a lot (resp. little) about past reviews. In the extreme case where $\eta = 1$, the rating is the difference between the accumulative number of all positive reviews and that of all negative ones; whereas if $\eta = 0$, the rating is such difference in the previous period.

In each period t, a continuum of infinitesimal customers arrive at the market. Each customer has a type V and requests at most one product. We assume that the willingness-to-pay of a new customer in period t is given by $V + \gamma(N_t)$. The customer type V captures her intrinsic valuation of the product, which is independent of the rating. The customer type V is uniformly positioned on the interval $(-\infty, \bar{V}_t]$ with density 1. Here, $\bar{V}_t$ can be interpreted as the maximum potential demand without online reviews. Clearly, there is an infinite mass of potential customers in the market, but, as we will show shortly, the actual demand in each period t is finite. We take this modeling approach to rule out potential corner solutions and thus simplify exposition. A similar approach has been used in the network effect literature (see Katz and Shapiro 1985). It is possible that a customer derives no intrinsic value from the product ($V \leq 0$). Such customers will not make a purchase without the online review system, but they may make a purchase in the presence of customer reviews. The function $\gamma(\cdot)$ captures how the online ratings impact customers, and is concavely increasing and twice continuously

[1] See, e.g., `https://tbfocus.com/blog/choose-good-seller-taobao`

[2] See, e.g., `https://econsultancy.com/blog/9366-ecommerce-consumer-reviews-why-you-need-them-and-how-to-use-them`

differentiable in the aggregate net rating N_t. Thus, the higher the rating of the firm, the more a customer wants to pay for the product. We normalize $\gamma(0)=0$, and assume that the marginal effect of rating diminishes as it goes to infinity (i.e., $\lim_{N_t\to+\infty}\gamma'(N_t)=0$). For technical tractability, we assume that customers are bounded rational so that they base their purchasing decisions on the current sales price and aggregate net rating, instead of rational expectations on future prices and ratings. Therefore, a type-V customer would make a purchase in period t if and only if $V+\gamma(N_t)\geq p_t$, where $p_t\in[\underline{p},\bar{p}]$ is the product price in period t and $\underline{p}$ (resp. $\bar{p}$) is the minimum (resp. maximum) allowable price. In each period t, there exists a random additive demand shock, ξ_t, which captures other uncertainties not explicitly modeled (e.g., the macro-economic condition of period t). Hence, the actual demand in period t is given by:

$$D_t(p_t,N_t):=\int_{-\infty}^{\bar{V}_t}1_{\{V+\gamma(N_t)\geq p_t\}}\,\mathrm{d}V+\xi_t=\bar{V}_t+\gamma(N_t)-p_t+\xi_t,$$

where ξ_t is independent of the price p_t and the rating N_t with $\mathbb{E}[\xi_t]=0$. Moreover, $\{\xi_t:t=T,T-1,\cdots,1\}$ are *i.i.d.* random variables with a continuous distribution. Without loss of generality, we assume that $\xi_t\geq-\bar{V}_t+\bar{p}$, which implies that $D_t(p_t,N_t)\geq0$ with probability 1, for all $p_t\in[\underline{p},\bar{p}]$ and N_t.

We now characterize the dynamics of the aggregate net ratings $\{N_t:t=T,T-1,\cdots,1\}$. As shown in (1), the rating of the next period N_{t-1} is determined by the new reviews of the current period and the exponentially smoothed rating of the past reviews. Hence, the aggregate net rating satisfies the following recursive pattern:

$$N_{t-1}=(r_t^+-r_t^-)+\eta N_t.$$

Customers who purchase the product in period t may post reviews in the online review system of the platform. When demand $D_t(p_t,N_t)$ exceeds the available inventory level x_t, some customers are backlogged and put onto a wait-list. Customers who get the products immediately have better consumer experiences and, thus, have a higher (resp. lower) chance to post a positive (resp. negative) review than a customer who is wait-listed. For each customer receiving the product immediately, the probability that she will post a positive (resp. negative) review is $\theta_1^+\in[0,1]$ (resp. $\theta_1^-\in[0,1]$). If a customer is wait-listed, the probability that she will post a positive (negative) review is $\theta_2^+\leq\theta_1^+$ (resp. $\theta_2^-\geq\theta_1^-$). Note that the default review on Taobao is a positive review[3]. This business practice motivates us to assume that customers are more likely to post a positive review regardless of whether stockout occurs. Hence, a customer is also more likely to post a positive review, even if she is wait-listed, i.e., $\theta_2^+\geq\theta_2^-$. Whether a customer will post an online review and whether she will post a positive review are independent of her type V, the rating N_t and the attribute of other customers.

[3] See, e.g., `https://tbfocus.com/blog/choose-good-seller-taobao`

This is because the rating a customer gives to the product/seller is determined primarily by the quality of the product itself, but not individual preferences and/or online reviews generated by other customers. We also assume there exists a random shock ϵ_t in the online rating dynamics, capturing any uncertainty not explicitly modeled. Therefore, given (x_t, p_t, N_t) in period t, there will be

$$r_t^+ = \theta_1^+(D_t(p_t, N_t) \wedge x_t) + \theta_2^+(D_t(p_t, N_t) - x_t)^+$$

positive reviews and

$$r_t^- = \theta_1^-(D_t(p_t, N_t) \wedge x_t) + \theta_2^-(D_t(p_t, N_t) - x_t)^+$$

negative reviews generated in this period. Therefore, the aggregate net rating at the beginning of the next period is given by

$$\begin{aligned} N_{t-1} &= (r_t^+ - r_t^-) + \eta N_t + \epsilon_t \\ &= (\theta_1^+ - \theta_1^-)(D_t(p_t, N_t) \wedge x_t) + (\theta_2^+ - \theta_2^-)(D_t(p_t, N_t) - x_t)^+ + \eta N_t + \epsilon_t \\ &= \theta(x_t \wedge D_t(N_t, p_t)) + (\theta - \sigma)(D_t(p_t, N_t) - x_t)^+ + \eta N_t + \epsilon_t \\ &= \theta D_t(p_t, N_t) - \sigma(D_t(p_t, N_t) - x_t)^+ + \eta N_t + \epsilon_t, \end{aligned} \tag{2}$$

where $\theta = \theta_1^+ - \theta_1^- > 0$, $\sigma = (\theta_1^+ - \theta_1^-) - (\theta_2^+ - \theta_2^-) \in [0, \theta]$, and ϵ_t is independent of the price p_t, the rating N_t, the available inventory x_t, and the demand perturbations $\{\xi_t : t = T, T-1, \cdots, 1\}$ with $\mathbb{E}[\epsilon_t] = 0$. Moreover, $\{\epsilon_t : t = T, T-1, \cdots, 1\}$ are *i.i.d.* random variables with a continuous distribution. We remark that the parameter σ measures the impact of service-level/inventory-availability on the aggregate net rating and, consequently, future demand. The larger the σ, the more significant the impact. Note that, when $\sigma = 0$, future demand depends on past *demand*; when $\sigma = \theta$, future demand depends on past *sales*.

The state of the inventory system is given by $(I_t, N_t) \in \mathbb{R} \times \mathbb{R}_+$, where

$I_t =$ the starting inventory level before replenishment in period t, $t = T, T-1, \cdots, 1$;

$N_t =$ the aggregate net rating of the seller at the beginning of period t, $t = T, T-1, \cdots, 1$.

The decisions of the firm is given by $(x_t, p_t) \in \hat{\mathcal{F}}(I_t) := [I_t, +\infty) \times [\underline{p}, \bar{p}]$, where

$x_t =$ the available inventory level after replenishment in period t, $t = T, T-1, \cdots, 1$;

$p_t =$ the sales price charged in period t, $t = T, T-1, \cdots, 1$.

In each period, the sequence of events unfolds as follows: At the beginning of period t, after observing the inventory level I_t and the rating N_t, the firm simultaneously chooses the inventory stocking level $x_t \geq I_t$ and the sales price p_t, and pays the ordering cost $c(x_t - I_t)$. The inventory procurement leadtime is assumed to be zero, so that the replenished inventory is received immediately. The demand $D_t(p_t, N_t)$ then realizes. The revenue, $p_t\mathbb{E}[D_t(p_t, N_t)]$, is collected. Unmet demand is fully backlogged. At the end of period t, the holding and backlogging costs are paid, the net inventory is carried over to the next period, and the aggregate net rating is updated according to the dynamics (2).

We introduce the following model primitives:

$$\alpha = \text{discount factor of revenues and costs in future periods, } 0 < \alpha < 1;$$

$$c = \text{inventory purchasing cost per unit ordered;}$$

$$b = \text{backlogging cost per unit backlogged at the end of a period;}$$

$$h = \text{holding cost per unit stocked at the end of a period.}$$

Without loss of generality, we make the following assumptions on the model primitives:

$b > (1-\alpha)c$: The backlogging penalty is higher than the saving from delaying an order to the next period, so that the firm will not backlog all of its demand;

$\underline{p} > b + \alpha c$: The margin for backlogged demand is positive.

For technical tractability, we make the following assumption throughout our theoretical analysis.

Assumption 1. *For each period t, $R_t(\cdot,\cdot)$ is jointly concave in $(p_t, N_t) \in [\underline{p}, \bar{p}] \times [0, +\infty)$, where*

$$R_t(p_t, N_t) := (p_t - b - \alpha c)(\bar{V}_t - p_t + \gamma(N_t)). \tag{3}$$

Given the sales price, p_t, and rating, N_t, of period t, $R_t(p_t, N_t)$ is the expected difference between the revenue and the total cost, which consists of ordering and backlogging costs, to satisfy the current demand in the next period. Hence, the joint concavity of $R_t(\cdot,\cdot)$ implies that such difference has decreasing marginal values with respect to the current sales price and rating. While the concavity of revenue with respect to price is a common assumption in the pricing literature, the joint concavity of $R_t(\cdot,\cdot)$ is a slightly stronger assumption, as it also captures the impact of aggregate net rating upon revenue, procurement cost, and backlogging cost. We remark that $R_t(\cdot, N_t)$ is strictly concave in p_t for any given N_t. Moreover, the monotonicity of $\gamma(\cdot)$ suggests that $R_t(\cdot,\cdot)$ is supermodular in (p_t, N_t).

Assumption 1 is essential to show the analytical results in this paper, because it ensures the concavity of the objective function in each period (see Lemma 3 in Appendix B). We characterize the necessary and sufficient conditions for this assumption in Appendix D. As shown by Lemma 7 in Appendix D, the necessary and sufficient condition for Assumption 1 is that $\underline{p} \geq \alpha c + b + \frac{M}{2}$, where $M := \sup\{-(\gamma'(N_t))^2/\gamma''(N_t) : N_t \geq 0\}$. Since the sensitivity of demand with respect to price $\frac{\partial \mathbb{E}[D_t(p_t,N_t)]}{\partial p_t} = -1$ is a constant, the condition $\underline{p} \geq \alpha c + b + \frac{M}{2}$ is equivalent to that the price elasticity of demand, $|\frac{\mathrm{d}\mathbb{E}[D_t(p_t,N_t)]/\mathbb{E}[D_t(p_t,N_t)]}{\mathrm{d}p_t/p_t}|$, is sufficiently high relative to the rating elasticity of demand,

$|\frac{\mathrm{d}\mathbb{E}[D_t(p_t,N_t)]/\mathbb{E}[D_t(p_t,N_t)]}{\mathrm{d}N_t/N_t}|$. Therefore, Assumption 1 has a clear and nonrestrictive economic interpretation: Compared with the primary demand leverage (i.e., sales price), the customer-generated rating has relatively less impact upon demand in general. In Appendix D.2, we also demonstrate that Assumption 1 can be satisfied for a wide variety of function families $\gamma(\cdot)$ (e.g., exponential, power, and logarithm functions), by giving some concrete examples of network externalities functions and deriving necessary and sufficient conditions for the concavity of $R_t(\cdot,\cdot)$.

4. Model Analysis

The purpose of this section is to simplify our model by demonstrating that the state space dimension of the dynamic program for the joint pricing and inventory replenishment problem can be reduced to 1. To this end, we first characterize the structure of the optimal policy in the presence of online customer reviews/ratings.

4.1. Optimal Policy

We now formulate the planning problem as a dynamic program. Define

$$v_t(I_t,N_t) := \text{the maximum expected discounted profits in periods } t, t-1, \cdots, 1, \text{ when starting period } t \text{ with an inventory level } I_t \text{ and online rating } N_t.$$

Without loss of generality, we assume that, in the last period (period 1), the excess inventory is salvaged with unit value c, and the backlogged demand is filled with ordering cost c, i.e., $v_0(I_0,N_0)=cI_0$ for any (I_0,N_0). We define $y_t(p_t,N_t) := \bar{V}_t - p_t + \gamma(N_t)$ as the expected demand with price p_t and online rating N_t. The optimal value function $v_t(I_t,N_t)$ satisfies the following recursive scheme:

$$v_t(I_t,N_t) = cI_t + \max_{(x_t,p_t)\in\hat{\mathcal{F}}(I_t)} J_t(x_t,p_t,N_t), \tag{4}$$

where $\hat{\mathcal{F}}(I_t) := [I_t,+\infty)\times[\underline{p},\bar{p}]$ denotes the set of feasible decisions and,

$$\begin{aligned}
J_t(x_t,p_t,N_t) &= R_t(p_t,N_t) + \beta x_t + \Lambda(x_t - y_t(p_t,N_t)) \\
&\quad + \mathbb{E}[\Psi_t(x_t - y_t(p_t,N_t) - \xi_t, \eta N_t + \theta(y_t(p_t,N_t)+\xi_t) - \sigma(y_t(p_t,N_t)+\xi_t - x_t)^+)], \\
\text{with } \Psi_t(x,y) &:= \alpha\mathbb{E}\{[v_{t-1}(x, y+\epsilon_t) - cx]\}, \\
\Lambda(x) &:= \mathbb{E}\{-(b+h)(x-\xi_t)^+\}, \\
\beta &:= b-(1-\alpha)c = \text{the effective monetary benefit of ordering one unit of inventory.}
\end{aligned}$$

The detailed derivation of $J_t(x_t,p_t,N_t)$ is given by (11) in Appendix B. Hence, for each period t, the firm selects

$$(x_t^*(I_t,N_t), p_t^*(I_t,N_t)) := \operatorname*{arg\,max}_{(x_t,p_t)\in\hat{\mathcal{F}}(I_t)} J_t(x_t,p_t,N_t) \tag{5}$$

as the optimal inventory and price policy contingent on the state variable (I_t, N_t).

As a stepping stone for our subsequent analysis, Lemma 3 in Appendix B shows that the value and objective functions $v_t(\cdot,\cdot)$, $J_t(\cdot,\cdot,\cdot)$, and $\Psi_t(\cdot,\cdot)$ are all jointly concave and continuously differentiable. In particular, the concavity and continuous differentiability of $J_t(\cdot,\cdot,\cdot)$ ensure that, the optimal price and inventory policy, $(x_t^*(I_t,N_t), p_t^*(I_t,N_t))$, is well-defined and can be obtained via first-order conditions. Moreover, we can define the inventory-independent optimizer in each period t as follows:

$$(x_t(N_t), p_t(N_t)) := \underset{x_t\in\mathbb{R}, p_t\in[\underline{p},\bar{p}]}{\arg\max}\ J_t(x_t,p_t,N_t). \tag{6}$$

In case of multiple optimizers, we select the lexicographically smallest one.

THEOREM 1. *For any t, the following statements hold:*

(a) If $I_t \le x_t(N_t)$, $(x_t^*(I_t,N_t), p_t^*(I_t,N_t)) = (x_t(N_t), p_t(N_t))$.

(b) If $I_t > x_t(N_t)$, $x_t^*(I_t,N_t) = I_t$ and $p_t^*(I_t,N_t) = \arg\max_{p_t\in[\underline{p},\bar{p}]} J_t(I_t,p_t,N_t)$.

(c) For any $I_t \in \mathbb{R}$ and $N_t \ge 0$, $x_t^*(I_t,N_t) > 0$.

Theorem 1 characterizes the optimal policy as a rating-dependent base-stock/list-price policy. If the starting inventory level I_t is below the rating-dependent base-stock level $x_t(N_t)$, it is optimal to order up to this base-stock level, and charge a rating-dependent list-price $p_t(N_t)$. If the starting inventory level is above the base-stock level, it is optimal to order nothing and charge an inventory-dependent sales price $p_t^*(I_t,N_t)$. Moreover, as shown in Theorem 1(c), the optimal period-t order-up-to level $x_t^*(I_t,N_t)$ is always positive for any inventory level I_t and aggregate net rating N_t.

4.2. State Space Dimension Reduction

The original dynamic program to characterize the optimal joint pricing and inventory policy has a state space of two dimensions (inventory level I_t and online rating N_t, see Section 4.1). Hence, it is analytically challenging and computationally complex to directly work with the recursive Bellman equation (4). Similar challenges have also been reported in the joint pricing and inventory management problem in the presence of reference price effects (Chen et al. 2016). In this subsection, we demonstrate that the dynamic program can actually be reduced to a much simpler one with a single-dimensional state space (rating N_t). Moreover, as long as the initial inventory level I_T is below the period-T optimal base-stock level $x_T(N_T)$, the optimal policy in each period t, $(x_t^*(I_t,N_t), p_t^*(I_t,N_t))$, is independent of the (stochastic and endogenous) inventory dynamics till this period $\{I_s : s = T, T-1, \cdots, t\}$ with probability 1. As we will show in Section 5, the state space dimension reduction serves as our stepping stone to deliver sharper insights on the managerial implications of online product reviews.

We first simplify the objective function $J_t(\cdot,\cdot,\cdot)$. Let $\Delta_t := x_t - y_t(p_t,N_t)$ be the safety-stock level in period t, given that the firm sets inventory stocking level x_t and price p_t, and the current rating of the seller is N_t. It is straightforward to show that, in period t, maximizing (x_t, p_t) over the feasible set $\hat{\mathcal{F}}(I_t)$ is equivalent to maximizing (Δ_t, p_t) over the feasible set $\mathcal{F}(I_t) := \{(\Delta_t, p_t) \in \mathbb{R} \times [\underline{p}, \bar{p}] : \Delta_t + y_t(p_t, N_t) \geq I_t\}$. Therefore, the objective function in period t can be written as:

$$O_t(\Delta_t, p_t, N_t) = Q_t(p_t, N_t) + \beta\Delta_t + \Lambda(\Delta_t) + \mathbb{E}[\Psi_t(\Delta_t - \xi_t, \eta N_t + \theta(y_t(p_t, N_t) + \xi_t) - \sigma(-\Delta_t + \xi_t)^+)],$$

where $Q_t(p_t, N_t) := R_t(p_t, N_t) + \beta y_t(p_t, N_t) = (p_t - c)(\bar{V}_t - p_t + \gamma(N_t))$ is jointly concave in (p_t, N_t). The detailed derivation of $O_t(\cdot,\cdot,\cdot)$ is given by (12) in Appendix B. For each rating N_t, we define $(\Delta_t(N_t), p_t(N_t))$ as the inventory-independent optimizer of $O_t(\cdot,\cdot,N_t)$, i.e.,

$$(\Delta_t(N_t), p_t(N_t)) := \underset{\Delta_t \in \mathbb{R}, p_t \in [\underline{p}, \bar{p}]}{\arg\max}\ O_t(\Delta_t, p_t, N_t). \tag{7}$$

Thus, $\Delta_t(N_t) = x_t(N_t) - y_t(p_t(N_t), N_t)$ is the optimal inventory-independent safety-stock level with online rating N_t.

We now employ sample path analysis to characterize the inventory dynamics under the optimal joint pricing and inventory policy.

LEMMA 1. *For each period t and any online rating N_t, we have*

$$\mathbb{P}[x_t(N_t) - D_t(p_t(N_t), N_t) \leq x_{t-1}(N_{t-1}) | N_t] = 1. \tag{8}$$

Lemma 1 is a key technical result of our paper. We show that, if the firm adopts the optimal policy and period-t starts with an inventory level below the optimal base-stock level (i.e., $I_t \leq x_t(N_t)$), the starting inventory level in the next period, period-$(t-1)$, $I_{t-1} = x_t(N_t) - D_t(p_t(N_t), N_t)$, will stay below the period-$(t-1)$ optimal base-stock level, $x_{t-1}(N_{t-1})$, with probability 1. The sample-path property (8) is a version of Condition 3(b) in Veinott (1965) in our joint pricing and inventory management model in the presence of online customer reviews. It has been widely shown in the inventory literature that this property (or its corresponding version in a different model) is essential in establishing the structural properties of an inventory system (e.g., Veinott 1965 and Section 6.3 of Porteus 2002).

An important implication of Lemma 1 is that once the starting inventory level falls below the optimal base-stock level in *some* period, it is optimal for the firm to replenish in *each* period thereafter throughout the planning horizon with probability 1. Our model works well for a new seller on the e-commerce platform, who has neither inventory nor reputation at the beginning of the sales season, i.e., $I_T = 0$. In this case, Theorem 1(c) and Lemma 1 together imply that, with probability 1, $I_t \leq x_t(N_t)$ for each period t.

Based on Lemma 1, we now show that the bivariate profit-to-go functions, $\{v_t(I_t, N_t) : t = T, T-1, \cdots, 1\}$, can be transformed into univariate ones of rating N_t by normalizing the value of inventory

cI_t. We construct the following dynamic program with a one-dimensional state space of online rating N_t:

$$\begin{aligned} \pi_t(N_t) &= \max_{\Delta_t \in \mathbb{R}, p_t \in [\underline{p}, \bar{p}]} O_t(\Delta_t, p_t, N_t), \text{ where} \\ O_t(\Delta_t, p_t, N_t) &= Q_t(p_t, N_t) + \beta\Delta_t + \Lambda(\Delta_t) + \mathbb{E}[G_t(\eta N_t + \theta(y_t(p_t, N_t) + \xi_t) - \sigma(-\Delta_t + \xi_t)^+)], \\ \text{with } G_t(y) &:= \alpha\mathbb{E}[\pi_{t-1}(y + \epsilon_t)], \text{ and } \pi_0(\cdot) \equiv 0. \end{aligned} \tag{9}$$

As shown in Lemma 4 in Appendix B, the solution to (9) forms the optimal rating-dependent safety-stock level and list-price in each period t, $(\Delta_t(N_t), p_t(N_t))$. Through reducing the original dynamic program (4) to the new one (9), we have essentially decoupled inventory and rating in the *state space*. To conclude this subsection, we give the following sharper characterization of the optimal joint pricing and inventory policy based on Theorem 1, Lemma 1, and Lemma 4.

Theorem 2. *Assume that $I_T \le x_T(N_T)$. In each period t, $(x_t^*(I_t, N_t), p_t^*(I_t, N_t)) = (x_t(N_t), p_t(N_t))$ with probability 1, where $x_t(N_t) = \Delta_t(N_t) + y_t(p_t(N_t), N_t)$. Moreover, $\{(\Delta_t(N_t), p_t(N_t)) : t = T, T-1, \cdots, 1\}$ is the solution to the Bellman equation* (9).

Theorem 2 shows that, as long as the planning horizon starts with an inventory level below the period-T optimal base-stock level (i.e., $I_T \le x_T(N_T)$), the optimal pricing and inventory policy in each period t, $(x_t^*(I_t, N_t), p_t^*(I_t, N_t))$, is identical to the optimal base-stock level and list-price, $(x_t(N_t), p_t(N_t))$, with probability 1. Although the firm holds inventory throughout the sales horizon, the optimal policy is independent of the inventory dynamics if the initial inventory level I_T is sufficiently low. As discussed above, in most applications, the firm holds zero initial inventory at the beginning of the sales season, i.e., $I_T = 0$. Hence, $(x_t^*(I_t, N_t), p_t^*(I_t, N_t)) = (x_t(N_t), p_t(N_t))$ for all (I_t, N_t) with probability 1. The state-space dimensionality reduction of the dynamic program also helps alleviate the complexity of numerically computing the optimal policy $\{(x_t^*(I_t, N_t), p_t^*(I_t, N_t)) : t = T, T-1, \cdots, 1\}$. As shown in Theorem 2, it suffices to compute the inventory-independent policy $\{(x_t(N_t), p_t(N_t)) : t = T, T-1, \cdots, 1\}$, which is the solution to a dynamic program with a single dimensional state space, (9). Based on Theorem 2, unless otherwise specified, we will confine our analysis to the properties of the optimal base-stock level and list-price $(x_t(N_t), p_t(N_t))$ for the rest of this paper.

5. Trade-off between Current Profits and Future Demands

This section strives to answer the following two questions: (a) How does the presence of online rating impact the operations decisions of the firm? (b) What strategies can the firm employ to exploit the network effect induced by online rating? The answers to these questions shed lights on the managerial implications of online customer reviews/ratings and the induced network effect. We show that the seller on an e-commerce platform with online ratings is facing the key trade-off between generating current profits and inducing future demands. This trade-off yields several interesting new insights on the operational implications of the online rating system.

5.1. Impact on Joint Pricing and Inventory Policy

We start the analysis with a comparison between our joint pricing and inventory management model with online product reviews and the benchmark model without (i.e., Federgruen and Heching 1999). The benchmark model corresponds to a special case of our model with $\gamma(\cdot) \equiv 0$.

Theorem 3. *Assume that two inventory systems are identical except that one with $\gamma(\cdot)$ and the other with $\hat{\gamma}(\cdot)$, where $\gamma(0) = \hat{\gamma}(0) = 0$ and $\hat{\gamma}(N_t) \geq \gamma(N_t) \equiv 0$ for all $N_t \geq 0$, i.e., the inventory system with $\gamma(\cdot)$ does not have online customer reviews. For each period t and any rating $N_t \geq 0$, the following statements hold:*

(a) $\hat{\Delta}_t(N_t) \geq \Delta_t(N_t) \equiv \Delta_*$, where $\Delta_* := \arg\max_\Delta\{\beta\Delta + \Lambda(\Delta)\}$ and the inequality is strict if $\sigma > 0$ and $\hat{\gamma}'(\cdot) > 0$;

(b) $\hat{x}_t(N_t) \geq x_t(N_t) \equiv x_*(t)$, where the inequality is strict if $\sigma > 0$ and $\hat{\gamma}'(\cdot) > 0$;

(c) there exists a threshold $\mathfrak{N}_t \geq 0$, such that $\hat{p}_t(N_t) \leq p_t(N_t) \equiv p_*(t)$ for $N_t \leq \mathfrak{N}_t$, whereas $\hat{p}_t(N_t) \geq p_t(N_t) \equiv p_*(t)$ for $N_t \geq \mathfrak{N}_t$.

With online product reviews, a service effect emerges: The firm should keep a higher safety-stock to induce higher potential ratings. As shown in Theorem 3(a,b), the presence of online reviews/ratings gives rise to the network effect, and drives the firm to increase the safety-stock and base-stock levels in each period t (i.e., $\hat{\Delta}_t(N_t) \geq \Delta_*$ and $\hat{x}_t(N_t) \geq x_*(t)$). Unlike the standard model without customer reviews, where the price only depends on inventory, the firm in the presence of the online review system should charge differentiated prices contingent on different ratings. Theorem 3(c) demonstrates an interesting effect of online rating upon the firm's pricing policy: The optimal price in the presence of online product reviews, $\hat{p}_t(\cdot)$, may be either higher or lower than that without, $p_*(t)$. More specifically, if the rating N_t is sufficiently low (i.e., below the threshold $\mathfrak{N}_t$), the optimal price in the presence of online reviews, $\hat{p}_t(\cdot)$, is lower than that without, $p_*(t)$. On the other hand, if the rating is sufficiently high (i.e., above the threshold $\mathfrak{N}_t$), the firm should increase the price in the presence of online reviews, i.e., $\hat{p}_t(N_t) \geq p_*(t)$. In the presence of customer-generated reviews, the firm faces the trade-off between decreasing the price to induce high future demands and increasing the price to exploit the current market. When the current rating is low ($N_t \leq \mathfrak{N}_t$), the firm should put higher weight on inducing future demands, so the optimal price is lower with online product reviews. Otherwise, when $N_t \geq \mathfrak{N}_t$, generating current profits outweighs inducing future demands, and, hence, the optimal price is higher in the presence of online reviews. In short, Theorem 3 reveals that, because of the trade-off between generating current profits and inducing future demands, the online rating system of e-commerce platforms gives rise to the service effect and rating-dependent pricing policy. We also remark that Theorem 3 and other comparative statics results of this paper have been established using the technique developed by Yang and Zhang (2022). Interested readers are referred to Appendices B and C for details.

We now proceed to investigate how the firm's optimal pricing and inventory policy responds to different ratings.

THEOREM 4. *For period t, assume that $\hat{N}_t > N_t$. We have: (a) $p_t(\hat{N}_t) \geq p_t(N_t)$; (b) $\Delta_t(\hat{N}_t) \leq \Delta_t(N_t)$; and (c) if $\gamma(\hat{N}_t) = \gamma(N_t)$, then $x_t(\hat{N}_t) \leq x_t(N_t)$.*

Theorem 4 sharpens our understanding of how the trade-off between generating current profits and inducing future demands impacts the pricing and inventory policy of the firm. More specifically, we show that the optimal price $p_t(N_t)$ is increasing in the current rating N_t, whereas the optimal safety-stock $\Delta_t(N_t)$ is decreasing in N_t. As the online rating increases (resp. decreases), the potential demand becomes larger (resp. smaller), and thus the firm is prompted to focus more on generating current profits (resp. inducing future demands) by increasing (resp. decreasing) the price. Analogously, with a higher current rating, the safety-stock level should be decreased. In this case, the service effect is weakened and the firm sets a lower safety-stock to save the procurement and holding costs. In summary, the firm puts more weight on generating current profits when the aggregate net rating is high, whereas it focuses more on inducing future demands when the rating is low.

Contrary to our intuition, Theorem 4(c) shows that the optimal base-stock level $x_t(N_t)$ may not necessarily be increasing in N_t. In the region where the firm is so reputable that the rating has no impact on demand (i.e., $\gamma'(\cdot) = 0$), the demand in the current period is not higher with a higher online rating, but an increased sales price (Theorem 4(a)) and a lower safety-stock level (Theorem 4(b)) reduce the resulting optimal base-stock level when the rating of the firm is higher.

The trade-off between generating current profits and inducing future demands gives rise to the service effect and rating-dependent pricing. Our next step is to study how these two phenomena evolve throughout the planning horizon. As shown in the following theorem, when the market is stationary (i.e., the highest customer type $\bar{V}_t$ is a constant with respect to the time index t), the presence of an online review system motivates the firm to set lower sales prices and higher safety-stock and base-stock levels at the beginning of the sales horizon.

THEOREM 5. *Assume that $\bar{V}_\tau = \bar{V}$ for all τ. For each $t \geq 2$ and any rating $N \geq 0$, we have (a) $\Delta_t(N) \geq \Delta_{t-1}(N)$, (b) $x_t(N) \geq x_{t-1}(N)$, and (c) $p_t(N) \leq p_{t-1}(N)$.*

Theorem 5 shows that the service effect becomes less intensive as the time approaches the end of the sales horizon. Specifically, with stationary customer type distribution (i.e., $\bar{V}_t$ is independent of time t) and the same online rating N (thus, the same potential demand), the optimal safety-stock $\Delta_t(N)$ and the optimal base-stock level $x_t(N)$ are decreasing, whereas the optimal sales price $p_t(N)$ is increasing over the sales horizon. In the presence of online product reviews, the firm should put more weight on inducing future demands at the beginning of the planning horizon, and turn to generating current profits as it approaches the end of the sales season. Hence, the firm offers discounts and increases the safety-stock (and, thus, base-stock) level to attract more customers to purchase the product and post their reviews on the e-commerce platform at the early stages of a sales season. On the other hand, the firm charges a higher price and sets a lower safety-stock to exploit the current market towards the end of the planning horizon.

Theorem 5 is consistent with the commonly used introductory price strategy under which price discounts are offered when a firm is launching its business on a platform. This strategy helps the firm accumulate online reviews and achieve a high rating, so that it can exploit the market and earn a high profit later. When the customer valuation is not stationary (i.e., $\bar{V}_t$ is time-dependent), the introductory price strategy may not necessarily be optimal. This is because, if the customer valuation is higher ($\bar{V}_t$ is larger) at the beginning of the sales season, the firm may charge a higher price to exploit the higher customer preference, as opposed to offering discounts.

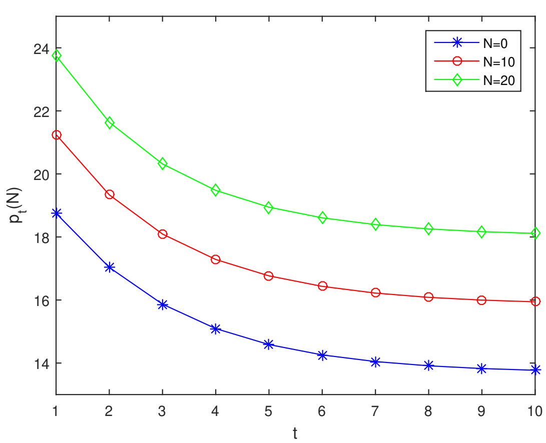

Figure 1 **Optimal List-Price $p_t(N)$**

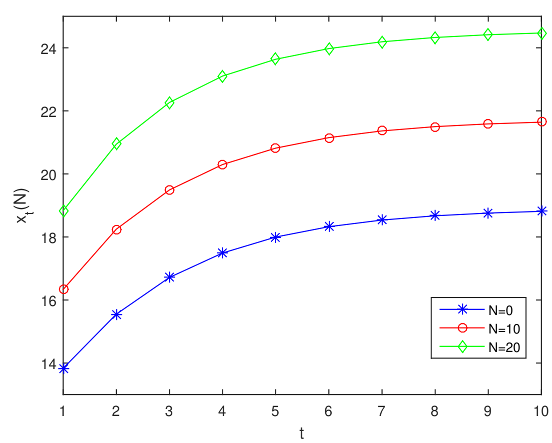

Figure 2 **Optimal Base-Stock Level $x_t(N)$**

To illustrate the behaviors of the optimal policy for different values of net rating throughout the planning horizon, we give a numerical example (the parameter specifications are provided in Appendix E). Figures 1-2 plot the optimal base-stock level and price $((x_t(N), p_t(N))$ for different t's and N's. As shown in Figures 1-2, these numerical illustrations further reinforce the theoretical predictions of Theorems 4 and 5, thus highlighting the essential trade-off between generating current profits and inducing future demands: (a) The firm should charge a higher price and stock more inventory if the net rating is higher; and (b) It is optimal to offer price discounts and stock more inventory at the beginning of the selling horizon.

To conclude this subsection, we analyze the impact of discount factor α on the optimal joint pricing and inventory policy. The discount factor measures how the firm values future profits relative to current profits. Studying its impact helps us better understand the trade-off between generating current profits and inducing future demands.

THEOREM 6. *Assume that two inventory systems are identical except that one with discount factor $\hat{\alpha}$, and the other with discount factor α, where $\hat{\alpha} > \alpha$. For each period t and any rating N_t, we have (a) $\hat{\Delta}_t(N_t) \geq \Delta_t(N_t)$; (b) $\hat{x}_t(N_t) \geq x_t(N_t)$; and (c) $\hat{p}_t(N_t) \leq p_t(N_t)$.*

Future profits are more valuable to the firm with a larger discount factor, so the firm puts more weight on inducing future demands relative to generating current profits. Therefore, as the discount factor α increases, the service effect gets strengthened (the safety-stock level $\Delta_t(N_t)$ and the base-stock level $x_t(N_t)$ both increase), and price discounts are offered (the price $p_t(N_t)$ decreases). Both changes facilitate the firm to attract more customers to post reviews and boost the rating, which helps better balance the trade-off between current profits and future demands.

5.2. Mean-Reverting Online Rating

We now proceed to study the stochastic evolution of aggregate net rating N_t. Interestingly, under the optimal pricing and inventory policy, the rating process $\{N_t : t = T, T-1, \cdots, 1\}$ follows a mean-reverting pattern.

THEOREM 7. *Assume that $\eta < 1$ and $I_T \leq x_T(N_T)$.*

(a) There exists a threshold value $\bar{N}_t \in (0, +\infty)$, such that $\mathbb{E}[N_{t-1}|N_t] > N_t$ for $N_t < \bar{N}_t$, and $\mathbb{E}[N_{t-1}|N_t] < N_t$ for $N_t > \bar{N}_t$.

(b) Assume that $\bar{V}_\tau \equiv \bar{V}$ for all τ and $T = +\infty$ (i.e., the infinite-horizon discounted reward criterion). We have N_t is ergodic and has a stationary distribution $\nu(\cdot)$, i.e., $\mathbb{P}(N_t \leq z) = \nu(z)$ for all t and all z. Moreover, there exists a threshold $\bar{N} = \lim_{t \to +\infty} \bar{N}_t \in (0, +\infty)$, such that, $\mathbb{E}[N_{t-1}|N_t] > N_t$ for $N_t < \bar{N}$, and $\mathbb{E}[N_{t-1}|N_t] < N_t$ for $N_t > \bar{N}$.

(c) $\bar{N}_t$ and, thus, $\bar{N}$ are increasing in η. In particular, as $\eta \uparrow 1$, $\bar{N}_t \uparrow +\infty$ and, thus, $\bar{N} \uparrow +\infty$.

As long as customers value recent reviews more than earlier ones (i.e., $\eta < 1$) and the initial inventory is below the initial base-stock level (i.e., $I_T \leq x_T(N_T)$), the rating dynamics exhibit a mean-reverting pattern: The future rating has an increasing trend in expectation ($\mathbb{E}[N_{t-1}|N_t] > N_t$) if the current rating N_t is below the "mean" $\bar{N}_t$, but would decrease in expectation ($\mathbb{E}[N_{t-1}|N_t] < N_t$) if N_t is above $\bar{N}_t$. Curiously, such mean-reversion persists even when customers value past reviews almost the same as current reviews (i.e., η is very close to 1). Theorem 7(b) further shows that, if the customer preference of the product is stationary (i.e., the maximum intrinsic customer valuation $\bar{V}_t$ is time-invariant), the rating process N_t has a stationary distribution $\nu(\cdot)$. Therefore, the "mean" $\bar{N}$ is time-invariant in this case. In Theorem 7(c), we demonstrate that if the discount factor for the aggregate net rating, η, is larger, the "mean" $\bar{N}_t$ (also $\bar{N}$) increases, and, thus, the product rating is more likely to grow. The mean-reverting pattern of the rating process clearly reflects the trade-off between generating current profits and inducing future demands via the network effect induced by online customer reviews. With a low current rating, the firm cares more about inducing future demands and, thus, strives to accumulate online reviews by adjusting its joint pricing and inventory decisions. On the other hand, with a high rating, the firm focuses on exploiting the current market and, hence, the future rating would fall in expectation. In particular, if the discount factor for the aggregate net rating, η, is larger, the current reviews are more influential on future customers. Hence, the trade-off between current profits and future demands is more intensive, and the firm adopts the joint pricing and inventory policy that drives the rating process to grow. In the limiting case where the product rating is cumulative (i.e., $\eta = 1$), the mean-reverting pattern is reduced to one in which the rating grows throughout the planning horizon with probability 1.

5.3. Dynamic Look-Ahead Heuristic

The goal of this subsection is to propose an easy-to-implement dynamic look-ahead heuristic policy, and quantitatively justify its effectiveness in the presence of online product reviews. This heuristic prescribes the joint pricing and inventory decisions that maximize the total profit of a (short) moving time-window throughout the planning horizon. We theoretically show that the profit gap between the optimal policy and the proposed heuristic decays exponentially as the moving time-window expands. Therefore, our heuristic can achieve small optimality gaps even with a short moving time-window. This is also verified by our numerical experiments. The key insight from our analysis is that the computationally efficient dynamic look-ahead heuristic policy could effectively leverage online customer-generated reviews by balancing generating current profits and inducing demands in the *near future*.

We first numerically examine the profit losses of the benchmark heuristic, the *myopic policy*. The myopic policy is the simplest heuristic policy and it completely ignores the future demand-inducing effect of online reviews. Adopting the myopic policy, the firm adjusts its price and inventory decisions to maximize the expected current-period profit. Hence, the myopic policy prescribes the solution to a newsvendor problem with endogenous pricing.

Throughout the numerical studies in this section, we assume that the maximum intrinsic valuation $\bar{V}_t = 30$ is stationary for each period t. The planning horizon length is $T = 20$. The function that captures the impact of the aggregate net rating on demand is $\gamma(N_t) = kN_t$ $(k \geq 0)$. The parameter

k measures the impact of product reviews on future demands. The larger the k, the higher the impact online reviews have on customers' purchasing decisions. Hence, the demand in each period t is $D_t(p_t, N_t) = 30 + kN_t - p_t + \xi_t$, where $\{\xi_t\}_{t=1}^T$ follow *i.i.d.* normal distributions with mean 0 and standard deviation 2 truncated so that $D_t(p_t, N_t) \geq 0$ with probability 1 for any (p_t, N_t). We set the discount factor $\alpha = 0.99$, the unit procurement cost $c = 8$, the unit holding cost $h = 1$, the unit backlogging cost $b = 10$, and the feasible price range $[\underline{p}, \bar{p}] = [0, 25]$. For simplicity, we assume that the random perturbation in the aggregate net rating dynamics, ϵ_t, is degenerate, i.e., $\epsilon_t = 0$ with probability 1. We also assume that a customer who gets the product immediately has the same probability to post a positive or negative review as that of a wait-listed customer, i.e., $\theta_1^+ = \theta_2^+$, $\theta_1^- = \theta_2^-$, and $\sigma = 0$. In the evaluation of the expected profits for the firm, we take $I_t = 0$ as the reference initial inventory level and $N_t = 0$ as the reference aggregate net rating. Our results are robust if we set a different initial inventory level and/or a different rating.

Let $v_t^0(I_t, N_t)$ be the expected total profits in periods $t, t-1, \cdots, 0$ under the myopic policy, if period t starts with inventory I_t and rating N_t. The metric of interest is

$$\lambda_m := \frac{v_t(\cdot,\cdot) - v_t^0(\cdot,\cdot)}{v_t(\cdot,\cdot)} \times 100\%, \text{ which evaluates the (relative) profit loss of the myopic policy.}$$

We report the numerical results with the parameters $t = 5, 10, 15, 20$, $k = 0.2, 0.5, 0.8$, $\theta = 0.2, 0.5, 0.8$, and $\eta = 0.2, 0.5, 0.8$.

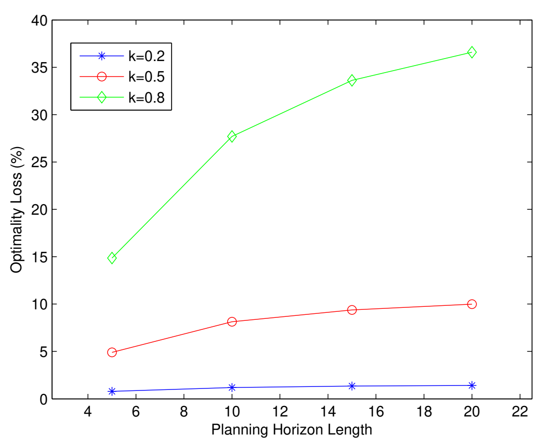

Figure 3 **Value of λ_m: $\theta = 0.5$, $\eta = 0.5$**

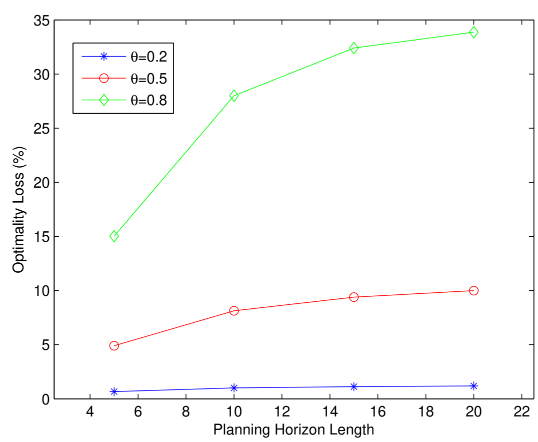

Figure 4 **Value of λ_m: $k = 0.5$, $\eta = 0.5$**

Figures 3-5 summarize the results of our numerical study on the profit performance of the myopic policy. As long as the impact of online rating on demand, k, the difference between the probability of a customer posting a positive review and that of a customer posting a negative review, θ, and the discount factor of aggregate net rating, η, are not too low (greater than 0.2 in our numerical cases), the myopic policy leads to a significant profit loss, which is at least 4.90% and can be as high as 36.60%. If k, θ, or η is large, the current operations decisions have great impact upon future customer-generated ratings, thus leading to intensive trade-off between generating current profits and inducing future demands. Therefore, the myopic policy results in significant profit losses if k, θ, and η are not too low. Another important implication of Figures 3-5 is that, if k, θ, and η are not too low, the profit loss of ignoring customer-generated reviews may be significant even when the planning horizon is short (i.e., $t=5$). This calls for caution that the seller on the e-commerce platform in the presence of customer reviews should not overlook the trade-off between generating current profits and inducing future demands even for a short sales horizon.

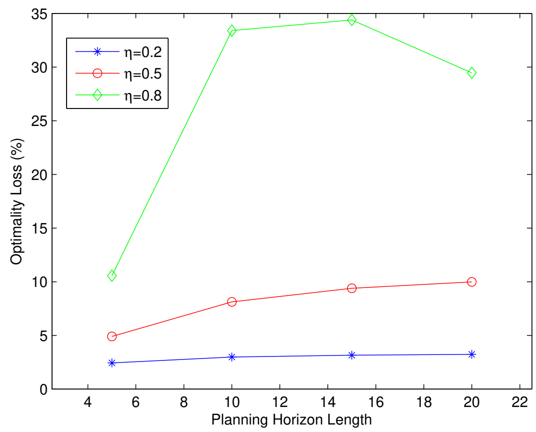

Figure 5 **Value of λ_m: $k=0.5$, $\theta=0.5$**

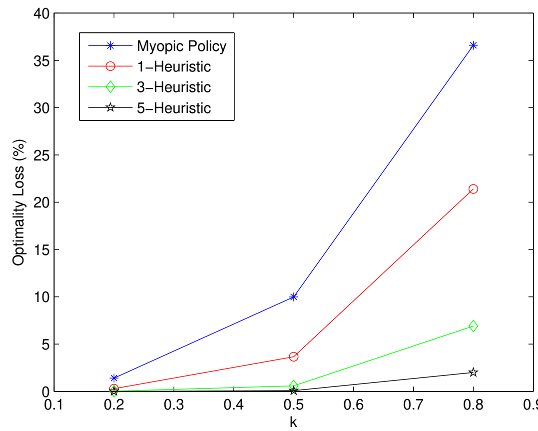

Figure 6 **Value of λ_m and λ_h^i: $\theta=0.5$, $\eta=0.5$**

The myopic policy completely ignores the demand-inducing effect of online ratings, and, therefore, gives rise to substantial profit losses. We now propose the dynamic look-ahead heuristic policy and study its value in the presence of customer-generated reviews/ratings. The key idea of the dynamic look-ahead heuristic is to (mildly) leverage the demand-inducing opportunities of online reviews/ratings while keeping the computational simplicity of the myopic policy. This heuristic policy balances generating current profits and inducing demands in the *near future* in the presence of online reviews. More specifically, in each period t, the firm adopts the joint pricing and inventory policy that maximizes the expected total discounted profits in the moving time-window of $w+1$ periods: The firm looks forward for w periods into the future, and maximizes the total profit from period t to period $\min\{t-w,1\}$. Similar dynamic look-ahead heuristics (also called the rolling-horizon procedures) are widely used in the literature to (approximately) solve complex dynamic programming problems with a high dimensional state space and a long planning horizon (see, e.g., Powell 2011). We refer to the dynamic look-ahead heuristic with the moving time-window of $w+1$ periods as the *w-heuristic* hereafter. Note that the 0-heuristic corresponds to the myopic policy, whereas the T-heuristic corresponds to the optimal policy. Obtaining the w-heuristic involves solving a dynamic program with planning horizon length $w+1$, so it is computationally efficient and easy to implement if w is small.

We first theoretically justify the effectiveness of the w-heuristic policy in the presence of online reviews. More specifically, we show that, if the customer preference of the product is stationary (i.e., $\bar{V}_t$ is independent of t), the gap between the optimal total profit and the total profit associated with the w-heuristic decays exponentially in the length of the moving time-window w. Formally, in the finite-horizon model (i.e., $T<+\infty$), let $v_t^w(I_t, N_t)$ be the expected profits in periods $t, t-1, \cdots, 0$ when the firm adopts the w-heuristic and period t starts with inventory I_t and online rating N_t. In the infinite-horizon discounted reward criterion model (i.e., $T=+\infty$), let $v^w(I,N)$ (resp. $v(I,N)$) be the expected discounted total profit when the firm adopts the w-heuristic (resp. optimal policy) and the planning horizon starts with inventory I and rating N.

Theorem 8. *Assume that $\bar{V}_\tau \equiv \bar{V}$ for all τ, $\eta<1$ and $I_T \le x_T(N_t)$.*

(a) If $T<+\infty$, we have $v_t^w(\cdot,\cdot) \le v_t^{w+1}(\cdot,\cdot) \le v_t(\cdot,\cdot)$ for all $w \ge 0$. Moreover, $v_t^w(\cdot,\cdot) = v_t(\cdot,\cdot)$ for $w \ge t-1$.

(b) If $T=+\infty$, we have $v^w(\cdot,\cdot) \le v^{w+1}(\cdot,\cdot) \le v(\cdot,\cdot)$ for all $w \ge 0$. There exist two constants $C>0$ and $\delta>0$, such that $\sup|v^w(\cdot,\cdot) - v(\cdot,\cdot)| \le C\mathrm{e}^{-\delta w}$. *Thus,* $\lim_{w\to+\infty} v^w(\cdot,\cdot) = v(\cdot,\cdot)$.

As shown in Theorem 8, the w-heuristic is sub-optimal, but its performance improves as the moving time-window length w increases. Thus, if the firm looks ahead more into the future, it can better balance current profits and future demands. Technically, this finding results from the monotonicity property (see Theorem 5) that, for a stationary market, the optimal price $p_t(\cdot)$ is decreasing, whereas the optimal safety-stock level $\Delta_t(\cdot)$ and base-stock level $x_t(\cdot)$ are increasing in the time index t. The choice of the moving time-window length w highlights the trade-off between computational efficiency and profitability in our model. A longer (resp. shorter) moving time-window yields higher (resp. lower) profits, but requires more (resp. less) computational effort as well. Interestingly, this trade-off is not very intensive in the sense that, the optimality gap of the w-heuristic, $\sup|v^w(\cdot,\cdot) - v(\cdot,\cdot)|$, decays exponentially in the forward-looking length w. Therefore, even with a short moving time-window and, thus, light computational burden, the w-heuristic could effectively exploit the network effect induced by online ratings, and achieve excellent profit performance. Specifically, to achieve an optimality gap of ϵ (i.e., $\sup|v^w(\cdot,\cdot) - v(\cdot,\cdot)| < \epsilon$), it suffices to employ the w-heuristic with moving time-window length $w \sim \mathcal{O}(\log(1/\epsilon))$.

We now proceed to numerically demonstrate the effectiveness of the w-heuristic in leveraging online ratings even for a short moving time-window. The metric of interest is

$$\lambda_h^w := \frac{v_t(\cdot,\cdot) - v_t^w(\cdot,\cdot)}{v_t(\cdot,\cdot)} \times 100\%, \text{ which measures the optimality gap of the } w\text{-heuristic.}$$

The numerical experiments are under the parameters $t = 20$, $k = 0.2, 0.5, 0.8$, $\theta = 0.2, 0.5, 0.8$, $\eta = 0.2, 0.5, 0.8$, and $w = 1, 3, 5$.

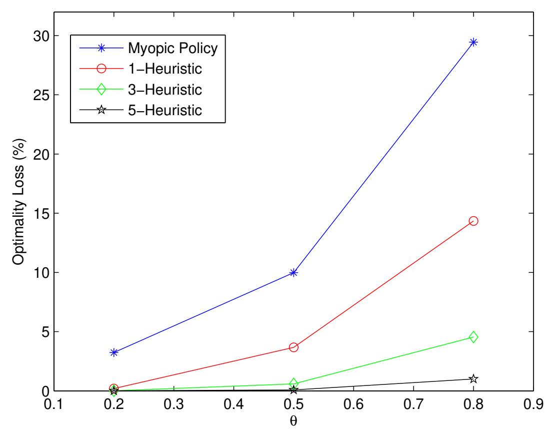

Figure 7 **Value of λ_m and λ_h^i: $k = 0.5$, $\eta = 0.5$**

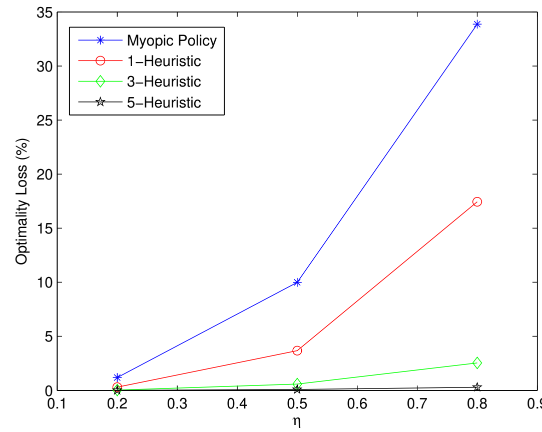

Figure 8 **Value of λ_m and λ_h^i: $k = 0.5$ $\theta = 0.5$**

Figures 6-8 summarize the results of our numerical study on the performance of w-heuristics. We find that, compared with the myopic policy that completely ignores the online reviews/ratings, the w-heuristics significantly improve the profits in the presence of customer-generated reviews. In particular, the 5-heuristic leads to substantially lower profit losses than those of the myopic policy (below 2%, in contrast to the above 30% optimality gap of the myopic policy). This confirms our theoretical prediction that, even for a small w ($w = 1, 3, 5$), the w-heuristic can achieve a very good profit performance. Therefore, the firm can effectively leverage the online product reviews/ratings by looking ahead into the near future and balancing the trade-off between generating current profits and inducing demands in the *near future*. Moreover, Figures 6-8 show that, as k, θ, or η increases, the trade-off between current profits and future demands becomes more intensive, and, thus, the look-ahead w-heuristics can deliver higher values to the firm when benchmarked against the myopic policy. We have also performed numerical analysis for the w-heuristics with longer moving time-windows (i.e, $w > 5$), which does not yield a significantly better performance over those with $w = 5$. This further demonstrates that, to exploit the network effect induced by online product reviews, it suffices to balance generating current profits and inducing demands in the *near future*. Finally, we remark that our numerical results are robust and continue to hold in the settings where the planning horizon length is greater than 20, and/or the market is non-stationary (i.e., the maximum intrinsic customer valuation $\bar{V}_t$ is time-dependent), and/or wait-listed customers have a different probability of posting a (positive or negative) review (i.e., $\sigma > 0$). For brevity, we only present the results for the case where $T = 20$, $\bar{V}_t$ is time-invariant, and $\sigma = 0$ in this paper.

In summary, the product review system of an e-commerce platform has several important operational implications upon the joint pricing and inventory policy of the firm. Most notably, online product reviews lead to a network effect, which creates another layer of complexity in balancing the trade-off between generating current profits and inducing future demands. Therefore, online reviews give rise to the service effect, rating-dependent pricing, and the mean-reverting pattern of the aggregate net rating process. Although completely ignoring the trade-off between current profits and future demands leads to substantial profit losses, it suffices to adopt the dynamic look-head heuristic policies that balance current profits and near-future demands. This family of heuristics are easy to implement, and achieve low optimality gaps with exponential decay in the length of the look-ahead time-window.

6. Extensions

In this section, we study two extensions of our base model: (a) the model with the paid-review strategy under which the firm provides monetary incentives for customers to leave reviews; and (b) the case where the starting inventory exceeds the optimal base-stock level in period T.

6.1. Paid Reviews

We study the operations implications to pay customers to leave reviews. Since the willingness-to-pay of the customers is increasing in the aggregate net rating, the firm may benefit from providing incentives to customers to leave more reviews for its product. The paid-review strategy is widely used to increase the number of reviews and the aggregate net rating[1].

For conciseness, we provide a macroscopic model which focuses on the total aggregate net rating increase in each period, which results from the firm's monetary incentive for customers to leave reviews. The micro-foundation of this model is given in Appendix F. Specifically, let c_n be the total cost the firm pays the customers in period t, which gives rise to an increase of net rating $n_t = n(c_n)$, where $n(\cdot)$ is a continuously differentiable and strictly increasing function with $n(0) = 0$. Define $c_n(n_t)$ as the total cost of increasing n_t aggregate net rating, where $c_n(\cdot)$ is the inverse of $n(\cdot)$. Furthermore, we assume that $c_n(\cdot)$ is convexly increasing. The convexity of $c_n(\cdot)$ captures the decreasing marginal value of paying customers to review, which is also justified by the micro-foundation of our model proposed in Appendix F. Note that although the paid reviews do not change the inventory dynamics of the firm, it does have some impact on the aggregate net rating dynamics. More specifically, by (2), the net rating at the beginning of period $t-1$ with paid reviews is given by:

$$N_{t-1} = \eta N_t + \theta D_t(p_t, N_t) - \sigma(D_t(p_t, N_t) - x_t)^+ + n_t + \epsilon_t.$$

[1] See, e.g., Green 2017, `https://www.kayako.com/blog/asking-for-reviews/`

We now formulate the dynamic program for the planning problem with the paid-review strategy. Define

$v_t^e(I_t, N_t) :=$ the maximum expected discounted profits with the paid-review strategy in periods $t, t-1, \cdots, 1, 0$, when starting period t with an inventory level I_t and aggregate net rating N_t;

and $(x_t^{e*}(I_t, N_t), p_t^{e*}(I_t, N_t), n_t^*(I_t, N_t))$ as the optimal inventory, pricing, and paid-review policy. As in the base model, we assume that, in the last period, the excess inventory is salvaged with unit value c, and the backlogged demand is filled with ordering cost c, i.e., $v_0^e(I_0, N_0) = cI_0$ for any (I_0, N_0). Employing similar dynamic programming and sample path analysis methods, Lemma 5 in Appendix B demonstrates that a rating-dependent base-stock/list-price/paid-review policy is optimal. The same sample path analysis technique as in the base model reduces the state space dimension of the dynamic program to 1, which further implies that as long as the initial inventory level I_T is below the optimal period-T base-stock level $x_T^e(N_T)$, the optimal policy is independent of the starting inventory level in each period with probability 1.

We remark that Theorems 3-8 are readily generalizable to the model with paid reviews. For brevity, these results are not presented in the paper, but available from the authors upon request. We now demonstrate the effectiveness of the paid-review strategy when the network effect of the online customer reviews is intensive.

THEOREM 9. *(a) Let* $0 < \iota < 1$, *and* $\bar{S}(N) := \sup\{z : \mathbb{P}(N_{t-1} \geq z | N_t = N) \geq \iota\}$. *If*

$$\alpha(1-\iota)(\underline{p} - c)\gamma'(\bar{S}(N)) > c_n'(0), \tag{10}$$

then $n_t^*(I_t, N) > 0$ *for all* I_t. *Moreover,* $\bar{S}(N)$ *is continuously increasing in* N *and, for each* $0 < \iota < 1$, *there exists an* $N_*(\iota) \geq 0$, *such that* (10) *holds for all* $N < N_*(\iota)$.

(b) If $\alpha(\sum_{\tau=1}^{t-1}(\alpha\eta)^{\tau-1})(\bar{p} - c)\gamma'(0) \leq c_n'(0)$, $n_t^*(I_t, N_t) \equiv 0$ *for all* I_t *and* $N_t \geq 0$.

Theorem 9 characterizes the dichotomy on when the firm should pay customers to review. Theorem 9(a) shows that, when the intensity of network effect for online rating is sufficiently strong (as characterized by inequality (10)), it is optimal for the firm to invest in paid reviews. In particular, the firm should adopt the network expansion strategy for a sufficiently low current network size (i.e., $n_t^*(I_t, N_t) > 0$ if $N_t \leq N_*(\iota)$). The intuition behind Theorem 9(a) is that, if a lower bound of the marginal value of paid reviews, $\alpha(1-\iota)(\underline{p} - c)\gamma'(\bar{S}(N))$, dominates its marginal cost $c_n'(0)$, the firm should provide monetary incentives for customers to review. Here, $\bar{S}(N)$ can be interpreted as the threshold such that, conditioned on $N_t = N$, the probability that the net rating in period $t-1$ exceeds $\bar{S}(N)$ is smaller than ι, regardless of the joint pricing and inventory policy the firm employs. On the other hand, Theorem 9(b) shows that if the network effect of customer reviews is not strong enough (i.e., $\alpha(\sum_{\tau=1}^{t-1}(\alpha\eta)^{\tau})(\bar{p} - c)\gamma'(0) \leq c_n'(0)$), it is optimal for the firm not to invest in paid reviews.

We now characterize the impact of the paid-review strategy upon the firm's optimal policy.

THEOREM 10. *Assume that two inventory systems are identical except that one with the network expansion strategy and the other without. For each period t and any network size $N_t \geq 0$, the following statements hold: (a) $\Delta_t^e(N_t) \leq \Delta_t(N_t)$; (b) $x_t^e(N_t) \leq x_t(N_t)$; (c) $p_t^e(N_t) \geq p_t(N_t)$; and (d) $\pi_t^e(N_t) \geq \pi_t(N_t)$, where the inequality is strict if $n_t(N_t) > 0$.*

Theorem 10 characterizes how the firm should adjust its joint pricing and inventory policy under the paid-review strategy. Since the sales price, the safety-stock, and the paid reviews all help induce future demands via the online customer rating, the monetary incentives for customers to provide reviews allow the firm to set a lower safety-stock and a higher sales price to generate higher profit in the current period. As a result, the optimal base-stock level is also lower in the presence of paid reviews. Theorem 10(d) further shows that, thanks to the network effect induced by online customer reviews, the paid-review strategy can improve the firm's profit.

In summary, the paid-review strategy helps the firm exploit network externalities of customer reviews by attracting more customers to provide reviews (with a cost) in each period. In particular, this strategy allows the firm to induce future demands with paid reviews, while generating higher current profit with a lower safety-stock and a higher sales price. The firm should invest in paid reviews when the network effect intensity (of aggregate online rating) is sufficiently strong.

6.2. Excess Starting Inventory

The main focus of our analysis is on the scenario with the starting inventory level below the optimal base-stock level ($I_t \leq x_t(N_t)$), because this scenario occurs with probability 1 as long as $I_T \leq x_T(N_T)$ (see Theorem 2). This section partially characterizes the structural properties of the optimal policy when the starting inventory exceeds the optimal base-stock level (i.e., $I_t > x_t(N_t)$).

THEOREM 11. *Assume that $\eta = 0$ and $\sigma = 0$. For each period t, the following statements hold:*

(a) $v_t(I_t, N_t)$ is supermodular in (I_t, N_t).

(b) $x_t^*(I_t, N_t)$ is continuously increasing in I_t and N_t.

(c) $p_t^*(I_t, N_t)$ is continuously decreasing in I_t, and continuously increasing in N_t.

(d) The optimal expected demand $y_t^*(I_t, N_t) := \bar{V}_t - p_t^*(I_t, N_t) + \gamma(N_t)$ is continuously increasing in I_t and N_t. Hence, $\mathbb{E}[N_{t-1}|N_t] = \theta y_t^*(I_t, N_t)$ is continuously increasing in I_t and N_t.

(e) The optimal safety-stock $\Delta_t^*(I_t, N_t) := x_t^*(I_t, N_t) - y_t^*(I_t, N_t)$ is continuously increasing in I_t and continuously decreasing in N_t.

In the special case where customers care about the most recent reviews only, and the probability of posting a positive or negative review is irrelevant to whether a customer is wait-listed (i.e., $\eta = 0$ and $\sigma = 0$), we are able to characterize some structural properties of the optimal policy for any starting inventory I_t. More specifically, Theorem 11(a) shows that the value function in each period

t, $v_t(I_t, N_t)$ is supermodular in (I_t, N_t). This is because, a higher rating leads to a larger potential demand and, thus, a higher marginal value of inventory. Analogously, the optimal expected demand $y_t^*(I_t, N_t)$ and the optimal expected rating in the next period are both increasing in the current rating N_t. As a consequence, if the current rating is higher, the firm increases the order-up-to level $x_t^*(I_t, N_t)$ to match demand with supply, and charges a higher sales price $p_t^*(I_t, N_t)$ to exploit the higher potential demand. Theorem 11 also shows how the starting inventory level I_t influences the optimal policy: A higher starting inventory level prompts the firm to increase the safety-stock and to decrease the sales price.

7. Concluding Remarks

This is the first paper in the literature to study the joint pricing and inventory management problem in the presence of online customer-generated reviews/ratings. We consider a seller on an e-commerce platform equipped with a customer review/rating system. A customer who purchases the product from the seller may leave a positive or a negative review, and the customers' willingness-to-pay is increasing in the aggregate net rating of the product, which is the difference between the (discounted) number of positive reviews and that of negative ones since the start of the planning horizon. Therefore, a network effect arises in the presence of online reviews, since potential customers are more likely to purchase if the product rating is higher. As a consequence, the firm faces an important trade-off between generating current profits and inducing future demands.

The optimal policy is a rating-dependent base-stock/list-price policy. Moreover, we demonstrate that, as long as the initial inventory level is below the initial optimal base-stock level, the inventory dynamics do not influence the optimal policy of the firm with probability 1. As a consequence, the state space dimension of the dynamic program can be reduced to one (aggregate net rating) by normalizing the current inventory value. Such state space dimension reduction greatly facilitates the analysis and computation of the optimal policy, and paves our way to deliver sharper insights from our model.

Our analysis highlights the key trade-off between generating current profits and inducing future demands through customer-generated reviews/ratings. The presence of online reviews gives rise to the service effect and rating-dependent pricing policy, both of which are absent without customer-generated reviews. Specifically, when the rating is low, the firm should decrease the sales price to exploit the demand-inducing effect of online reviews. Otherwise, when the rating is high, the firm should increase the sales price to leverage the high potential demand. From the inter-temporal perspective, the firm should put more weight on inducing future demands at early stages of a sales season than at later stages. Thus, when the market is stationary, the firm employs the introductory price strategy that offers early purchase discounts to induce high demands and boost ratings at the beginning of the sales season. As a consequence of the firm's effort to balance the trade-off between current profits and future demands, the online rating process follows an interesting mean-reverting pattern: If the current rating is low (resp. high), it has an increasing (resp. decreasing) trend in expectation. We also find that the dynamic look-ahead heuristic that maximizes the total profits in a (short) moving time-window can achieve small optimality gaps which decay exponentially in the length of the moving time-window. Therefore, it suffices to balance generating current profits and inducing demands in the near future. The commonly adopted paid-review strategy facilitates the retailer to (partially) separate generating current profits and inducing future demands via the network effect of online reviews.

There are several avenues for extending this paper. For example, whereas we focus the back-order model for tractability, lost-sales are also very common in the e-commerce setting. Addressing the technical challenge to establish the joint concavity for the joint pricing and inventory control model in the presence of customer reviews and lost-sales (e.g., Feng et al. 2020) will be a fruitful future research direction. Furthermore, it will be worthwhile studying the price and inventory competition in the presence of online customer reviews. For a large e-commerce platform like Taobao, there are millions of sellers competing with each other, all facing a network effect induced by platform-facilitated online customer reviews. The analysis of the equilibrium pricing and inventory policy in this competitive landscape would offer insights for operating a store on a large-scale e-commerce platform. Second, the focus of our paper is on how the seller could optimize his pricing and inventory policy. It will be interesting to examine, given the seller's policy, how the platform should design its online review system to maximize the gross merchant value (GMV), which is a key metric of an e-commerce platform.

Acknowledgments

The authors thank the Department Editor Qi Annabelle Feng, the anonymous associate editor, and two referees for their very helpful and constructive comments, which have led to significant improvements in both the content and exposition of this paper. The second author is grateful for the financial support from the National Natural Science Foundation of China [NSFC71802133] and the Shanghai Eastern Scholar Program [QD2018053].

References

[1] Allon, G., D. J. Zhang. 2017. Managing service systems in the presence of social networks. *Working paper.*

[2] Aviv, Y. 2002. Gaining benefits from joint forecasting and replenishment processes: The case of auto-correlated demand. *Manufacturing & Service Operations Management.* **4**(1), 55-74.

[3] Bensaid, B., J.P. Lesne. 1996. Dynamic monopoly pricing with network externalities. *International Journal of Industrial Organization* **14**, 837-855.

[4] Bernstein, F., Y. Li, K. Shang. 2016. A simple heuristic for joint inventory and pricing models with lead time and backorders. *Management Science*, **62**(8), 2358-2373.

[5] Besbes, O., M. Scarsini. 2016. On information distortions in online ratings. *Operations Research*, **66**(3), 597-610.

[6] Cabral, L., D. Salant, G. Woroch. 1999. Monopoly pricing with network externalities. *International Journal of Industrial Organization*, **17**, 199-214.

[7] Chao, X., X. Gong, S. Zheng. 2014. Optimal pricing and inventory policies with reliable and random-yield suppliers: characterization and comparison. *Annals of Operations Research*, **241**(1-2), 35-51.

[8] Chen, X., P. Hu, S. Shum, Y. Zhang. 2016. Dynamic stochastic inventory management with reference price effects. *Operations Research*, **64**(6), 1529-1536.

[9] Chen, X., Z. Pang, L. Pan. 2014. Coordinating inventory control and pricing strategies for perishable products. *Operations Research*, **62**(2), 284-300.

[10] Chen, X., M. Sim, D. Simchi-Levi, P. Sun. 2007. Risk aversion in inventory management. *Operations Research*, **55**(5), 828-842.

[11] Chen, X., D. Simchi-Levi. 2004a. Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The finite horizon case. *Operations Research*, **52**(6), 887-896.

[12] Chen, X., D. Simchi-Levi. 2004b. Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The infinite horizon case. *Mathematics of Operations Research*, **29**(3), 698-723.

[13] Chen, X., D. Simchi-Levi. 2006. Coordinating inventory control and pricing strategies with random demand and fixed ordering cost: The continuous review model. *Operations Research Letters*, **34**, 323-332.

[14] Chen, X., D. Simchi-Levi. 2012. Pricing and inventory management. R. Philips O. Ozalp, eds. *The Oxford Handbook of Pricing Management.* Oxford University Press, USA, 784-824.

[15] Chen, Y., S. Ray, Y. Song. 2006. Optimal inventory and pricing control policy in periodic review system with fixed ordering cost and lost-sales. *Naval Research Logistics*, **53**(2), 117-136.

[16] Crapis, D., B. Ifrach, C. Maglaras, M. Scarsini. 2017. Monopoly pricing in the presence of social learning. *Management Science*, **63**(11), 3586-3608.

[17] Dellarocas, C. 2003. The digitization of word of mouth: Promise and challenges of online feedback mechanisms. *Management Science*, **49**(10), 1407–1424.

[18] Dellarocas, C., Zhang, X., Awad, N. F. 2007. Exploring the value of online product reviews in forecasting sales: The case of motion pictures. *Journal of Interactive Marketing*, **21**(4), 23–45.

[19] Dhebar, A., S. Oren. 1986. Dynamic nonlinear pricing in networks with interdependent demands. *Operations Research*, **34**(3), 384-394.

[20] Federgruen, A., A. Heching. 1999. Combined pricing and inventory control under uncertainty. *Operations Research*, **47**(3), 454-475.

[21] Feng, Q. 2010. Integrating dynamic pricing and replenishment decisions under supply capacity uncertainty. *Management Science*, **56**(12), 2154-2172.

[22] Feng, Q., Luo, S., Shanthikumar, J. G. 2020. Integrating dynamic pricing with inventory decisions under lost sales. *Management Sci.*, **66**(5), 2232-2247.

[23] Feng, Y., M. Hu. 2017. Blockbuster or niche? Competitive strategy under network effects. *Working paper.*

[24] Gong, X., X. Chao, S. Zheng. 2014. Dynamic pricing and inventory management with dual suppliers of different lead times and disruption risks. *Production Operations Management*, **23**(12), 2058-2074.

[25] Graves, S. 1999. A single-item inventory model for a nonstationary demand process. *Manufacturing & Service Operations Management*, **1**(1), 50-61.

[26] Hu, M., Wang, Z., Feng, Y. 2020. Information disclosure and pricing policies for sales of network goods. *Operations Research*, **68**(4), 1162-1177.

[27] Hu, M., J. Milner, J. Wu. 2016. Liking and following the newsvendor: Operations and marketing policies under social influence. *Management Science*, **62**(3), 867-879.

[28] Huh, W., G. Janakiraman. 2008. (s, S) optimality in joint inventory-pricing control: An alternative approach. *Operations Research*, **56**(3), 783-790.

[29] Johnson, G. D., H. E. Thompson. 1975. Optimality of myopic inventory policies for certain dependent demand processes. *Management Science*, **21**(11), 1303-1307.

[30] Katz, M., C. Shapiro. 1985. Network externalities, competition, and compatibility. *American Economics Review*, **75**(3), 424-440.

[31] Katz, M., C. Shapiro. 1986. Technology adoption in the presence of network externalities. *Journal of Political Economics*, **94**(4), 822-841.

[32] Li, Q., S. Zheng. 2006. Joint inventory replenish and pricing control for systems with uncertain yield and demand. *Operations Research*, **54**(4), 696-705.

[33] Mai, Y., B. Hu, S. Pekec. 2018. Courteous or crude? Managing user conduct to improve on-demand service platform performance. *Working paper.*

[34] Pang, Z., F. Y. Chen, Y. Feng. 2012. A note on the structure of joint inventory-pricing control with leadtimes. *Operations Research*, **60**(3), 581-587.

[35] Papanastasiou, Y., N. Savva. 2017. Dynamic pricing in the presence of social learning and strategic consumers. *Management Science*, **63**(4), 919-939.

[36] Petruzzi, N.C., M. Dada. 1999. Pricing and the newsvendor problem: A review with extensions. *Operations Research*, **47**(2), 183-194.

[37] Petruzzi, N.C., M. Dada. 2002. Dynamic pricing and inventory control with learning. *Naval Research Logistics*, **49**(3), 303-325.

[38] Porteus, E. 2002. *Foundations of Stochastic Inventory Theory*. Stanford Business Books, Stanford.

[39] Powell, W. 2011. *Approximate Dynamic Programming: Solving the Curses of Dimensionality, 2nd Edition*. Wiley. Hoboken.

[40] Shin, D, A. Zeevi. 2017. Dynamic pricing and learning with online product reviews. *Working paper*.

[41] Shou, B., R. Zheng. 2017. Pricing in social networks with strategic consumers. *Working paper*.

[42] Veinott, A. 1965. Optimal policy for a multi-product, dynamic, nonstationary inventory problem. *Management Science*, **12**(3), 206-222.

[43] Wang, R., Z. Wang. 2017. Consumer choice models with endogenous network effects. *Management Science*, **63**(11), 3531-3997.

[44] Xiao, G., N. Yang, R. Zhang. 2015. Dynamic pricing and inventory management under fluctuating procurement costs. *Manufacturing & Service Operations Management*, **17**(3), 321-334.

[45] Yang, J. 2012. Inventory and price control under time-consistent coherent and Markov risk measure. *Working paper*.

[46] Yang, N., R. Zhang. 2014. Dynamic pricing and inventory management under inventory-dependent demand. *Operations Research*, **62**(5), 1077-1094.

[47] Yang, N., R. Zhang. 2022. Comparative statics analysis of inventory management model with dynamic pricing, market environment fluctuation, and delayed differentiation. *Production and Operations Management*, **31**(1), 341-357.

[48] Yang, Y, Y. Chen, Y. Zhou. 2014. Coordinating inventory control and pricing strategies under batch ordering. *Operations Research*, **62**(1), 25–37.

[49] Yu, M., L. Debo. R. Kapuscinski. 2015. Strategic waiting for consumer-generated quality information: Dynamic pricing of new experience goods. *Management Science*, **62**(2), 410-435.

[50] Zhou, S.X., X. Chao. 2014. Dynamic pricing and inventory management with regular and expedited supplies. *Production and Operations Management*, **23**(1), 65-80.

Online Appendices to "Dynamic Pricing and Inventory Management in the Presence of Online Reviews"

Nan Yang and Renyu Zhang

We use ∂ to denote the derivative operator of a single variable function, ∂_x to denote the partial derivative operator of a multi-variable function with respect to variable x, and $1_{\{\cdot\}}$ to denote the indicator function. For any multivariate continuously differentiable function $f(x_1, x_2, \cdots, x_n)$ and $\tilde{x} := (\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n)$ in $f(\cdot)$'s domain, $\forall i$, we use $\partial_{x_i} f(\tilde{x}_1, \tilde{x}_2, \cdots, \tilde{x}_n)$ to denote $\partial_{x_i} f(x_1, x_2, \cdots, x_n)|_{x=\tilde{x}}$. We use $\epsilon_1 \overset{d}{=} \epsilon_2$ to denote that two random variables ϵ_1 and ϵ_2 follow the same distribution.

Appendix A: Table of Notations

Table 1 Summary of Notations

T :	planning horizon length	t:	period-index, labeled backwards
N_t :	aggregate net rating of period t	η :	discount factor for reviews
V :	intrinsic valuation of the customers	$\bar{V}_t$:	maximum potential demand without reviews
$\gamma(\cdot)$:	impact of online ratings on demand	p_t:	price of period t
$D_t(\cdot,\cdot)$:	demand function of the product in period t	θ:	net rating contribution ratio of demand
σ :	impact of inventory availability on net rating	α:	discount factor of the firm
c:	inventory purchasing cost	b:	backlogging cost of inventory
h:	holding cost of inventory	$R_t(\cdot,\cdot)$:	adjusted revenue of period t
β:	effective benefit of ordering inventory	I_t :	inventory level of period t
x_t:	base-stock level of period t	Δ_t :	safety-stock of period t
ξ_t:	demand perturbation in period t	ϵ_t:	perturbation of net rating in period t
n_t:	net rating increase with paid reviews in period t	$y_t(\cdot,\cdot) = \mathbb{E}[D_t(\cdot,\cdot)]$:	expected demand in period t
$c_n(n_t)$:	cost of generating n_t additional net rating	r_t^+ (resp. r_t^-):	number of positive (resp. negative) reviews

Appendix B: Auxiliary Results

We first give some auxiliary results serving as building blocks of our subsequent analysis. The proofs of these results are given in Appendix C. The following lemma (adapted from Lemma 4 in Yang and Zhang 2022) is essential in proving various comparative statics results in our model. This comparative statics analysis technique has been proposed by Yang and Zhang (2022).

LEMMA 2. *Let $F_i(z, Z)$ be a continuously differentiable and jointly concave function in (z, Z) for $i = 1, 2$, where $z \in [\underline{z}, \bar{z}]$ ($\underline{z}$ and $\bar{z}$ might be infinite) and $Z \in \mathbb{R}^n$. For $i = 1, 2$, let $(z_i, Z_i) := \arg\max_{(z,Z)} F_i(z, Z)$ be the optimizers of $F_i(\cdot,\cdot)$. If $z_1 < z_2$, we have: $\partial_z F_1(z_1, Z_1) \leq \partial_z F_2(z_2, Z_2)$.*

Next, we develop the preliminary concavity and differentiability properties of the value and objective functions in the following lemma, which serves as a stepping stone for our subsequent analysis.

LEMMA 3. *For each $t = T, T-1, \cdots, 1$, the following statements hold:*

(a) $\Psi_t(\cdot,\cdot)$ is jointly concave and continuously differentiable in (x, y). Moreover, $\Psi_t(x, y)$ is decreasing in x and increasing in y.

(b) $J_t(\cdot,\cdot,\cdot)$ is jointly concave and continuously differentiable in (x_t, p_t, N_t).

(b) $v_t(\cdot,\cdot)$ is jointly concave and continuously differentiable in (I_t, N_t). Moreover, $v_t(I_t, N_t)$ is increasing in N_t, and $v_t(I_t, N_t) - cI_t$ is decreasing in I_t.

Lemma 3 proves that, in each period t, the objective and value functions are concave and continuously differentiable. Moreover, after normalized with the value of inventory, the profit-to-go, $v_t(I_t, N_t) - cI_t$, is decreasing in the inventory level I_t and increasing in the network size N_t. Results similar to lemma 3 have also been established in other joint pricing and inventory management settings (see, e.g., Theorem 1 in Federgruen and Heching 1999).

In the following, we give the derivation for the objective functions in each period t.

Derivation of $J_t(x_t, p_t, N_t)$:

$$
\begin{aligned}
J_t(x_t,p_t,N_t) &= -cI_t + \mathbb{E}\{p_t D_t(p_t,N_t) - c(x_t - I_t) - h(x_t - D_t(p_t,N_t))^+ - b(x_t - D_t(p_t,N_t))^- \\
&\quad + \alpha v_{t-1}(x_t - D_t(p_t,N_t), \eta N_t + \theta D_t(p_t,N_t) - \sigma(D_t(p_t,N_t) - x_t)^+ + \epsilon_t) | N_t\}, \\
&= (p_t - \alpha c - b) y_t(p_t,N_t) + (b - (1-\alpha)c)x_t + \mathbb{E}\{-(h+b)(x_t - y_t(p_t,N_t) - \xi_t)^+ \\
&\quad + \alpha[v_{t-1}(x_t - y_t(p_t,N_t) - \xi_t, \theta(y_t(p_t,N_t) + \xi_t) + \eta N_t - \sigma(y_t(p_t,N_t) + \xi_t - x_t)^+ + \epsilon_t) \\
&\quad - c(x_t - y_t(p_t,N_t) - \xi_t)] | N_t\} \\
&= R_t(p_t,N_t) + \beta x_t + \Lambda(x_t - y_t(p_t,N_t)) + \\
&\quad + \mathbb{E}[\Psi_t(x_t - y_t(p_t,N_t) - \xi_t, \eta N_t + \theta(y_t(p_t,N_t) + \xi_t) - \sigma(y_t(p_t,N_t) + \xi_t - x_t)^+)]. \qquad (11)
\end{aligned}
$$

Derivation of $O_t(\cdot,\cdot,\cdot)$:

$$
\begin{aligned}
O_t(\Delta_t,p_t,N_t) :=& J_t(\Delta_t + y_t(p_t,N_t), p_t, N_t) \\
=& R_t(p_t,N_t) + \beta(\Delta_t + y_t(p_t,N_t)) + \Lambda(\Delta_t) \\
& + \mathbb{E}[\Psi_t(\Delta_t - \xi_t, \eta N_t + \theta(y_t(p_t,N_t) + \xi_t) - \sigma(-\Delta_t + \xi_t)^+)] \\
=& Q_t(p_t,N_t) + \beta\Delta_t + \Lambda(\Delta_t) + \mathbb{E}[\Psi_t(\Delta_t - \xi_t, \eta N_t + \theta(y_t(p_t,N_t) + \xi_t) - \sigma(-\Delta_t + \xi_t)^+)].
\end{aligned} \qquad (12)
$$

The following lemma paves our way to reduce the original dynamic program (4), which has a two-dimension state space, to one with a single-dimension state space.

Lemma 4. *The sequence of functions $\{\pi_t(\cdot) : t = T, T-1, \cdots, 1\}$ satisfy the following statements:*

(i) $\pi_t(\cdot)$ is concavely increasing and continuously differentiable in N_t with $\pi_t(N_t) := \max\{O_t(\Delta_t, p_t, N_t) : \Delta_t \in \mathbb{R}, p_t \in [\underline{p}, \bar{p}]\}$ for any $N_t \geq 0$;

(ii) $v_t(I_t, N_t) = cI_t + \pi_t(N_t)$ for all $N_t \geq 0$ and $I_t \leq x_t(N_t)$;

(iii) For all $N_t \geq 0$ and $\Delta_t \leq \Delta_t(N_t)$,

$$
O_t(\Delta_t,p_t,N_t) = Q_t(p_t,N_t) + \beta\Delta_t + \Lambda(\Delta_t) + \mathbb{E}[G_t(\eta N_t + \theta(y_t(p_t,N_t) + \xi_t) - \sigma(-\Delta_t + \xi_t)^+)], \qquad (13)
$$

where $G_t(y) := \alpha\mathbb{E}[\pi_{t-1}(y + \epsilon_t)]$;

(iv) $(\Delta_t(N_t), p_t(N_t))$ maximizes the right-hand side of equation (13) over the feasible set $\mathbb{R} \times [\underline{p}, \bar{p}]$.

More specifically, it follows immediately from Lemma 4 that the optimal rating-dependent safety-stock level and list-price in each period t, $(\Delta_t(N_t), p_t(N_t))$, can be recursively determined by solving the dynamic program (9).

For the model with the paid-review strategy (Section 6.1), we have the following lemma that characterizes the optimal policy in the presence of paid reviews.

LEMMA 5. *Iteratively define a sequence of functions* $\{\pi_t^e(N_t): t = T, T-1, \cdots, 1\}$ *and a sequence of pricing and inventory policies* $\{(x_t^e(N_t), p_t^e(N_t), n_t(N_t)): t = T, T-1, \cdots, 1\}$ *as follows:*

$$\begin{aligned} \pi_t^e(N_t) &= \max_{(\Delta_t, p_t, n_t) \in \mathcal{F}_e} O_t^e(\Delta_t, p_t, n_t, N_t), \qquad (14) \\ \textit{where } O_t^e(x_t, p_t, n_t, N_t) &= Q_t(p_t, N_t) + \beta\Delta_t + \Lambda(\Delta_t) - c_n(n_t) \\ &\quad + \mathbb{E}[G_t^e(\eta N_t + \theta(y_t(p_t, N_t) + \xi_t) - \sigma(-\Delta_t + \xi_t)^+ + n_t)], \\ \textit{with } G_t^e(y) &:= \alpha\mathbb{E}[\pi_{t-1}^e(y + \epsilon_t)], \ \pi_0^e(\cdot) \equiv 0, \\ (\Delta_t^e(N_t), p_t^e(N_t), n_t(N_t)) &:= \operatorname*{arg\,max}_{(\Delta_t, p_t, n_t) \in \mathcal{F}_e} O_t^e(x_t, p_t, n_t, N_t), \ \textit{and } \mathcal{F}_e := \mathbb{R} \times [\underline{p}, \bar{p}] \times \mathbb{R}_+. \end{aligned}$$

Also define $x_t^e(N_t) := \Delta_t^e(N_t) + y_t(p_t^e(N_t), N_t)$.

(a) If $I_t \le x_t^e(N_t)$, $(x_t^{e*}(I_t, N_t), p_t^{e*}(I_t, N_t), n_t^*(I_t, N_t)) = (x_t^e(N_t), p_t^e(N_t), n_t(N_t))$ *and* $v_t^e(I_t, N_t) = cI_t + \pi_t^e(N_t)$. *Otherwise,* $I_t > x_t^e(N_t)$, $x_t^{e*}(I_t, N_t) = I_t$.

(b) If $I_T \le x_T^e(N_T)$, $(x_t^{e*}(I_t, N_t), p_t^{e*}(I_t, N_t), n_t^*(I_t, N_t)) = (x_t^e(N_t), p_t^e(N_t), n_t(N_t))$ *for all* t *with probability 1.*

Appendix C: Proofs of Statements

For completeness, we also give the proof of Lemma 2, which is identical to Lemma 4 in Yang and Zhang (2022).

Proof of Lemma 2: Since $z_1 < z_2$, we have $\underline{z} \le z_1 < z_2 \le \bar{z}$. Hence, $\partial_z F_1(z_1, Z_1) \begin{cases} = 0 & \text{if } z_1 > \underline{z}, \\ \le 0 & \text{if } z_1 = \underline{z}; \end{cases}$ and $\partial_z F_2(z_2, Z_2) \begin{cases} = 0 & \text{if } z_2 < \bar{z}, \\ \ge 0 & \text{if } z_2 = \bar{z}, \end{cases}$ i.e., $\partial_z F_1(z_1, Z_1) \le 0 \le \partial_z F_2(z_2, Z_2)$. *Q.E.D.*

Proof of Lemma 3: We prove parts (a) - (c) together by backward induction.

We first show, by backward induction that if $v_{t-1}(I_{t-1}, N_{t-1}) - cI_{t-1}$ is jointly concave in (I_{t-1}, N_{t-1}), decreasing in I_{t-1}, and increasing in N_{t-1}, (i) $\Psi_t(\cdot,\cdot)$ is jointly concave in (x, y), decreasing in x, and increasing in y; (ii) $J_t(\cdot,\cdot,\cdot)$ is jointly concave in (x_t, p_t, N_t); and (iii) $v_t(I_t, N_t) - cI_t$ is jointly concave in (I_t, N_t), decreasing in I_t, and increasing in N_t. It is clear that $v_0(I_0, N_0) - cI_0 = 0$ is jointly concave, decreasing in I_0, and increasing in N_0. Hence, the initial condition holds.

Assume that $v_{t-1}(I_{t-1}, N_{t-1}) - cI_{t-1}$ is jointly concave in (I_{t-1}, N_{t-1}), decreasing in I_{t-1}, and increasing in N_{t-1}. Since concavity and monotonicity are preserved under expectation, $\Psi_t(\cdot,\cdot)$ is jointly concave in (x, y), decreasing in x, and increasing in y. Analogously, $\Lambda(x)$ is concavely decreasing in x. We now verify that

$$\Phi_t(x_t, p_t, N_t) := \mathbb{E}[\Psi_t(x_t - \bar{V}_t + p_t - \gamma(N_t) - \xi_t, \eta N_t + \theta(\bar{V}_t - p_t + \gamma(N_t) + \xi_t) - \sigma(\bar{V}_t - p_t + \gamma(N_t) + \xi_t - x_t)^+)]$$

is jointly concave in (x_t, p_t, N_t) and increasing in N_t. Since $\gamma(\cdot)$ is increasing in N_t, $\sigma \le \theta$, and $\Psi_t(x, y)$ is decreasing in x and increasing in y, $\Phi_t(x_t, p_t, N_t)$ is increasing in N_t. Let $\lambda \in [0, 1]$, $x_* = \lambda x_t + (1-\lambda)\hat{x}_t$, $p_* = \lambda p_t + (1-\lambda)\hat{p}_t$, and $N_* = \lambda N_t + (1-\lambda)\hat{N}_t$, we have:

$$\begin{aligned} &\lambda\Phi_t(x_t, p_t, N_t) + (1-\lambda)\Phi_t(\hat{x}_t, \hat{p}_t, \hat{N}_t) \qquad (15) \\ = \ &\lambda\mathbb{E}[\Psi_t(x_t - \bar{V}_t + p_t - \gamma(N_t) - \xi_t, \theta(\bar{V}_t - p_t + \gamma(N_t) + \xi_t) - \sigma(\bar{V}_t - p_t + \gamma(N_t) + \xi_t - x_t)^+ + \eta N_t)] \\ &+ (1-\lambda)\mathbb{E}[\Psi_t(\hat{x}_t - \bar{V}_t + \hat{p}_t - \gamma(\hat{N}_t) - \xi_t, \theta(\bar{V}_t - \hat{p}_t + \gamma(\hat{N}_t) + \xi_t) - \sigma(\bar{V}_t - \hat{p}_t + \gamma(\hat{N}_t) + \xi_t - \hat{x}_t + \xi_t)^+ + \eta\hat{N}_t)] \\ \le \ &\mathbb{E}[\Psi_t(A_t - \xi_t, \eta N_* + \theta B_t + \sigma C_t)] \end{aligned}$$

where

$$
\begin{aligned}
A_t =& x_* - \bar{V}_t + p_* - \lambda\gamma(N_t) - (1-\lambda)\gamma(\hat{N}_t),\\
B_t =& \lambda(\bar{V}_t - p_t + \gamma(N_t) + \xi_t) + (1-\lambda)(\bar{V}_t - \hat{p}_t + (1-\lambda)\gamma(\hat{N}_t) + \xi_t),\\
C_t =& -\lambda(\bar{V}_t - p_t + \gamma(N_t) + \xi_t - x_t)^+ - (1-\lambda)(\bar{V}_t - \hat{p}_t + \gamma(\hat{N}_t) + \xi_t - \hat{x}_t)^+,
\end{aligned}
$$

and the inequality follows from the joint concavity of $\Psi_t(\cdot,\cdot)$. Since $\gamma(\cdot)$ is concave, $A_t \geq x_* - \bar{V}_t + p_* - \gamma(N_*)$. Since $\gamma(\cdot)$ is concave, $\theta \geq \sigma$, and $-(\cdot)^+$ is concavely decreasing, $\theta B_t + \sigma C_t \leq \theta(\bar{V}_t - p_* + \gamma(N_*) + \xi_t) - \sigma(\bar{V}_t - p_* + \gamma(N_*) + \xi_t - x_*)^+$. Therefore, since $\Psi_t(x,y)$ is decreasing in x and increasing in y,

$$
\begin{aligned}
&\mathbb{E}[\Psi_t(A_t - \xi_t, \eta N_* + \theta B_t + \sigma C_t)]\\
\leq& \mathbb{E}[\Psi_t(x_* - \bar{V}_t + p_* - \gamma(N_*) - \xi_t, \eta N_* + \theta(\bar{V}_t - p_* + \gamma(N_*) + \xi_t) - \sigma(\bar{V}_t - p_* + \gamma(N_*) + \xi_t - x_*)^+)],
\end{aligned}
$$

i.e., $\Psi_t(\cdot,\cdot,\cdot)$ is jointly concave in (x_t, p_t, N_t).

Since $\Lambda(x) = \mathbb{E}\{-(h+b)(x-\xi_t)^+\}$ is concavely decreasing in x, similar argument to the case of $\Phi_t(x_t,p_t,N_t)$ implies that $\Lambda(x_t - \bar{V}_t + p_t - \gamma(N_t))$ is jointly concave in (x_t,p_t,N_t) and increasing in N_t. By Assumption 1, $R_t(p_t,N_t)$ is jointly concave in (p_t,N_t). Moreover, since $\gamma(\cdot)$ is increasing in N_t, $R_t(p_t,N_t)$ is increasing in N_t as well. Hence,

$$
\begin{aligned}
J_t(x_t,p_t,N_t) =& R_t(p_t,N_t) + \beta x_t + \Lambda(x_t - \bar{V}_t + p_t - \gamma(N_t))\\
&+ \mathbb{E}[\Psi_t(x_t - \bar{V}_t + p_t - \gamma(N_t) - \xi_t, \eta N_t + \theta(\bar{V}_t - p_t + \gamma(N_t) + \xi_t) - \sigma(\bar{V}_t - p_t + \gamma(N_t) + \xi_t - x_t)^+)]
\end{aligned}
$$

is jointly concave in (x_t,p_t,N_t) and increasing in N_t.

Since concavity is preserved under maximization (e.g., Boyd and Vanderberghe 2004), the joint concavity of $v_t(\cdot,\cdot)$ follows directly from that of $J_t(\cdot,\cdot,\cdot)$. Note that for any $\hat{I}_t > I_t$, $\hat{\mathcal{F}}(\hat{I}_t) \subseteq \hat{\mathcal{F}}(I_t)$. Thus,

$$
\begin{aligned}
v_t(\hat{I}_t,N_t) - c\hat{I}_t =& \max_{(x_t,p_t)\in\hat{\mathcal{F}}(\hat{I}_t)} J_t(x_t,p_t,N_t)\\
\leq& \max_{(x_t,p_t)\in\hat{\mathcal{F}}(I_t)} J_t(x_t,p_t,N_t)\\
=& v_t(I_t,N_t) - cI_t.
\end{aligned}
$$

Hence, $v_t(I_t,N_t) - cI_t$ is decreasing in I_t. Since $J_t(x_t,p_t,N_t)$ is increasing in N_t for any (x_t,p_t,N_t), for any $\hat{N}_t > N_t$,

$$
\begin{aligned}
v_t(I_t,\hat{N}_t) - cI_t =& \max_{(x_t,p_t)\in\hat{\mathcal{F}}(I_t)} J_t(x_t,p_t,\hat{N}_t)\\
\geq& \max_{(x_t,p_t)\in\hat{\mathcal{F}}(I_t)} J_t(x_t,p_t,N_t)\\
=& v_t(I_t,N_t) - cI_t.
\end{aligned}
$$

Thus, $v_t(I_t,N_t) - cI_t$ is increasing in N_t.

Second, we show, by backward induction, that if $v_{t-1}(\cdot,\cdot)$ is continuously differentiable, $\Psi_t(\cdot,\cdot)$, $J_t(\cdot,\cdot,\cdot)$, and $v_t(\cdot,\cdot)$ are continuously differentiable as well. For $t=0$, $v_0(I_0,N_0) = cI_0$ is clearly continuously differentiable. Thus, the initial condition holds.

If $v_{t-1}(\cdot,\cdot)$ is continuously differentiable, $\Psi_t(\cdot,\cdot)$ is continuously differentiable with partial derivatives given by

$$
\partial_x \Psi_t(x,y) = \mathbb{E}\{\alpha[\partial_{I_t} v_{t-1}(x, y+\epsilon_t) - c]\}, \tag{16}
$$

$$
\partial_y \Psi_t(x,y) = \alpha\mathbb{E}\{\partial_{N_{t-1}} v_{t-1}(x, y+\epsilon_t)\}, \tag{17}
$$

where the exchangeability of differentiation and expectation is easily justified using the canonical argument (see, e.g., Theorem A.5.1 in Durrett 2010, the condition of which can be easily verified observing the continuity of the partial derivatives of $v_{t-1}(\cdot,\cdot)$, and that the distributions of ξ_t and ϵ_t are continuous.). Moreover, since ξ_t is continuously distributed, $\Lambda(\cdot)$ and $\Phi_t(x_t, p_t, N_t)$ are continuously differentiable with

$$\begin{aligned}
\partial_{x_t}\Phi_t(x_t,p_t,N_t) =& \mathbb{E}[\partial_x\Psi_t(x_t-y_t(p_t,N_t)-\xi_t, \theta(y_t(p_t,N_t)+\xi_t)-\sigma(y_t(p_t,N_t)+\xi_t-x_t)^+ + \eta N_t)] \\
& + \sigma\mathbb{E}[\partial_y\Psi_t(x_t-y_t(p_t,N_t)-\xi_t, \theta(y_t(p_t,N_t)+\xi_t)-\sigma(y_t(p_t,N_t)+\xi_t-x_t)^+ + \eta N_t)1_{\{\xi_t\geq x_t-y_t(p_t,N_t)\}}] \\
\partial_{p_t}\Phi_t(x_t,p_t,N_t) =& \mathbb{E}[\partial_x\Psi_t(x_t-y_t(p_t,N_t)-\xi_t, \theta(y_t(p_t,N_t)+\xi_t)-\sigma(y_t(p_t,N_t)+\xi_t-x_t)^+ + \eta N_t)] \\
& - \theta\mathbb{E}[\partial_y\Psi_t(x_t-y_t(p_t,N_t)-\xi_t, \theta(y_t(p_t,N_t)+\xi_t)-\sigma(y_t(p_t,N_t)+\xi_t-x_t)^+ + \eta N_t)] \\
& + \sigma\mathbb{E}[\partial_y\Psi_t(x_t-y_t(p_t,N_t)-\xi_t, \theta(y_t(p_t,N_t)+\xi_t)-\sigma(y_t(p_t,N_t)+\xi_t-x_t)^+ + \eta N_t)1_{\{\xi_t\geq x_t-y_t(p_t,N_t)\}}]
\end{aligned}$$

By the continuous differentiability of $\gamma(\cdot)$, $R_t(\cdot,\cdot)$ is continuously differentiable. Therefore, $J_t(\cdot,\cdot,\cdot)$ is continuously differentiable in (x_t, p_t, N_t). If $I_t \neq x_t(N_t)$, the continuous differentiability of $v_t(\cdot,\cdot)$ follows immediately from that of $J_t(\cdot,\cdot,\cdot)$ and the envelope theorem. To complete the proof, it suffices to check that, for all $N_t \geq 0$, the left and right partial derivatives of the first variable at $(x_t(N_t), N_t)$, $\partial_{I_t}v_t(x_t(N_t)-, N_t)$ and $\partial_{I_t}v_t(x_t(N_t)+, N_t)$ are equal. By the envelope theorem,

$$\begin{cases}\partial_{I_t}v_t(x_t(N_t)-,N_t) = & c, \\ \partial_{I_t}v_t(x_t(N_t)+,N_t) = & c+\beta+\partial_x\Lambda(x_t(N_t)-y_t(p_t(N_t),N_t))+\partial_x\Phi_t(x_t(N_t),p_t(N_t),N_t).\end{cases}$$

The first-order condition with respect to x_t implies that

$$\beta+\partial_x\Lambda(x_t(N_t)-y_t(p_t(N_t),N_t))+\partial_x\Phi_t(x_t(N_t),p_t(N_t),N_t)=0.$$

Therefore, $\partial_{I_t}v_t(x_t(N_t)-,N_t)=\partial_{I_t}v_t(x_t(N_t)+,N_t)=c$. This completes the induction and, thus, the proof of Lemma 3. *Q.E.D.*

Proof of Theorem 1: Parts (a)-(b) follow immediately from the joint concavity of $J_t(\cdot,\cdot,N_t)$ in (x_t,p_t) for any $N_t \geq 0$.

We now show **part (c)** by backward induction. More specifically, we prove that if $x_{t-1}(N_{t-1})>0$ for all $N_{t-1}\geq 0$, $x_t(N_t)>0$ for all $N_t\geq 0$. Since $v_0(I_0,N_0)=cI_0$, $\Psi_1(x,y)\equiv 0$. Since $D_1\geq 0$ with probability 1, $\partial_x\Lambda(-\bar{V}_1+p_1-\gamma(N_1))=0$ for all $p_1\in[\underline{p},\bar{p}]$ and $N_1\geq 0$. Hence, for any $p_1\in[\underline{p},\bar{p}]$ and $N_1\geq 0$,

$$\partial_{x_1}J_1(0,p_1,N_1)=\beta-\partial_x\Lambda(-\bar{V}_1+p_1-\gamma(N_1))=\beta>0.$$

Hence, $x_1(N_1)>0$ for any $N_1\geq 0$. Thus, the initial condition is satisfied.

Now we assume that $x_{t-1}(N_{t-1})>0$ for all $N_{t-1}\geq 0$ and $x_t(\tilde{N}_t)\leq 0$ for some $\tilde{N}_t\geq 0$. Thus,

$$I_{t-1}=x_t(\tilde{N}_t)-D_t(p_t(\tilde{N}_t),\tilde{N}_t)\leq 0<x_{t-1}(\tilde{N}_{t-1})$$

almost surely, where

$$\tilde{N}_{t-1}=\eta\tilde{N}_t+\theta D_t(p_t(\tilde{N}_t),\tilde{N}_t)-\sigma(D_t(p_t(\tilde{N}_t),N_t)-x_t(\tilde{N}_t))^+ +\epsilon_t.$$

Thus, by **part (a)**, $\partial_{I_{t-1}} v_{t-1}(I_{t-1}, \tilde{N}_{t-1}) = c$ almost surely, when conditioned on $N_t = \tilde{N}_t$. Hence, conditioned on $N_t = \tilde{N}_t$, $\partial_x \Psi_t(x,y) = \alpha \mathbb{E}\{\partial_{I_{t-1}} v_{t-1}(I_{t-1}, \tilde{N}_{t-1}) - c | N_t = \tilde{N}_t\} = c - c = 0$, when (x_t, p_t) lies in the neighborhood of $(x_t(\tilde{N}_t), p_t(\tilde{N}_t))$. Since $x_t(\tilde{N}_t) \le 0$, $\partial_x \Lambda(x_t(\tilde{N}_t) - \bar{V}_t + p_t - \gamma(\tilde{N}_t)) = 0$ for all $p_t \in [\underline{p}, \bar{p}]$. Hence, for any $p_t \in [\underline{p}, \bar{p}]$,

$$\partial_{x_t} J_t(x_t(\tilde{N}_t), p_t, \tilde{N}_t) = \beta - \partial_x \Lambda(x_t(\tilde{N}_t) - \bar{V}_t + p_t - \gamma(\tilde{N}_t)) = \beta > 0.$$

Hence, $x_t(\tilde{N}_t) > 0$, which contradicts the assumption that $x_t(\tilde{N}_t) \le 0$ is the optimizer of (6) when $N_t = \tilde{N}_t$. Therefore, $x_t(N_t) > 0$ for all $N_t \ge 0$, whenever $x_t(\tilde{N}_t) > 0$ for all $\tilde{N}_t$. This completes the induction and, thus, the proof of **part (c)**. *Q.E.D.*

Proof of Lemma 1: We first show that (8) holds for the case $\sigma = 0$. Specifically, with backward induction, we show that, if $\sigma = 0$, (i) (8) holds for each period t and (ii) $\Delta_t(N_t) = \Delta_*$ for all t and N_t, where $\Delta_* = \arg\max_\Delta \{\beta\Delta + \Lambda(\Delta)\}$. When $t = 1$, $\Psi_t(\cdot,\cdot) \equiv 0$, optimizing the objective function in period 1, (12), indicates that $\Delta_1(N_1) = \Delta_*$ for all N_1 and (8) automatically holds.

We now show that if (i) and (ii) hold for period $t-1$, they also hold for period t. First, we prove that $\Delta_t(N_t) \le \Delta_*$. If, to the contrary, $\Delta_t(N_t) > \Delta_*$, Lemma 2 yields that

$$\begin{aligned} &\partial_{\Delta_t}[Q_t(p_t(N_t), N_t) + \beta\Delta_t(N_t) + \Lambda(\Delta_t(N_t)) + \mathbb{E}[\Psi_t(\Delta_t(N_t) - \xi_t, \theta y_t(p_t(N_t), N_t) + \eta N_t + \theta\xi_t)] \\ &\ge \partial_\Delta[\beta\Delta + \Lambda(\Delta)], \end{aligned}$$

i.e.,

$$\beta + \Lambda'(\Delta_t(N_t)) + \mathbb{E}[\partial_x \Psi_t(\Delta_t(N_t) - \xi_t, \theta y_t(p_t(N_t), N_t) + \eta N_t + \theta\xi_t)] \ge \beta + \Lambda'(\Delta_*).$$

The concavity of $\Lambda(\cdot)$ implies that $\Lambda'(\Delta_t(N_t)) \le \Lambda'(\Delta_*)$. Moreover, since $\Psi_t(x,y)$ is decreasing in x, $\mathbb{E}[\partial_x \Psi_t(\Delta_t(N_t) - \xi_t, \theta y_t(p_t(N_t), N_t) + \eta N_t + \theta\xi_t)] \le 0$. Therefore, $\Lambda'(\Delta_t(N_t)) = \Lambda'(\Delta_*)$ and $\mathbb{E}[\partial_x \Psi_t(\Delta_t(N_t) - \xi_t, \theta y_t(p_t(N_t), N_t) + \eta N_t + \theta\xi_t)] = 0$. Thus, by the first-order condition with respect to Δ_t, $(p_t(N_t), \Delta_*)$ is also the optimal price and safety-stock level, which is strictly lexicographically smaller than $(p_t(N_t), \Delta_t(N_t))$. This contradicts the assumption that we select the lexicographically smallest optimizer in each period. Hence, $\Delta_t(N_t) \le \Delta_*$ for all $N_t \ge 0$.

We now show that (8) holds. Note that, conditioned on N_t,

$$\begin{aligned} x_t(N_t) - D_t(p_t(N_t), N_t) - x_{t-1}(N_{t-1}) &= \Delta_t(N_t) - \xi_t - x_{t-1}(N_{t-1}) \\ &\le \Delta_* - \xi_t - x_{t-1}(N_{t-1}) \\ &= -y_{t-1}(p_{t-1}(N_{t-1}), N_{t-1}) - \xi_t \\ &= y_{t-1}(\bar{p}, 0) - \xi_t \\ &\overset{d}{=} y_{t-1}(\bar{p}, 0) - \xi_{t-1} \\ &= -D_{t-1}(\bar{p}, 0) \end{aligned}$$

where the first inequality follows from $\Delta_t(N_t) \le \Delta_*$, the second equality from the inductive hypothesis that $x_{t-1}(N_{t-1}) = y_{t-1}(p_{t-1}(N_{t-1}), N_{t-1}) + \Delta_*$ for all $N_{t-1} \ge 0$, the second inequality from $y_{t-1}(p_t, N_t) \ge$

$y_{t-1}(\bar{p},0)$, and the third equality from from $\xi_{t-1}\stackrel{d}{=}\xi_t$. Because $D_{t-1}(\bar{p},0)\geq 0$ with probability 1, conditioned on N_t, we have

$$\mathbb{P}[x_t(N_t)-D_t(p_t(N_t),N_t)\leq x_{t-1}(N_{t-1})|N_t]=\mathbb{P}[-D_{t-1}(p_{t-1}(N_{t-1}),N_{t-1})\leq 0|N_t]=1,$$

i.e., conditioned on N_t, $x_t(N_t)-D_t(p_t(N_t),N_t)\leq x_{t-1}(N_{t-1})$ with probability 1, i.e., (8) holds for period t.

To complete the induction, we show that $\Delta_t(N_t)=\Delta_*$. If, to the contrary $\Delta_t(N_t)<\Delta_*$, Lemma 2 implies that

$$\begin{aligned}&\partial_{\Delta_t}[Q_t(p_t(N_t),N_t)+\beta\Delta_t(N_t)+\Lambda(\Delta_t(N_t))+\mathbb{E}[\Psi_t(\Delta_t(N_t)-\xi_t,\theta y_t(p_t(N_t),N_t)+\eta N_t+\theta\xi_t)]\\ \leq\ &\partial_{\Delta}[\beta\Delta+\Lambda(\Delta)].\end{aligned}$$

On the other hand, (8) implies that $\mathbb{E}[\partial_x\Psi_t(\Delta_t(N_t)-\xi_t,\theta y_t(p_t(N_t),N_t)+\eta N_t+\theta\xi_t)]=0$. Thus,

$$\beta+\Lambda'(\Delta_t(N_t))\leq\beta+\Lambda'(\Delta_*).$$

The concavity of $\Lambda(\cdot)$ indicates that $\Lambda'(\Delta_t(N_t))=\Lambda'(\Delta_*)$. Since Δ_* is the smallest minimizer of $[\beta\Delta+\Lambda(\Delta)]$, we have $\Delta_t(N_t)=\Delta_*$. This completes the induction and, thus, the proof of the sample path property (8) for the case with $\sigma=0$.

To complete the proof, it suffices to show (8) holds for the case $0<\sigma\leq\theta$. Observe that the above argument continues to hold if $\xi_t<\Delta_t(N_t)$, since, in this case, the inventory stocking level does not affect future network size evolutions. By Theorem 1(c), if $\xi_t\geq\Delta_t(N_t)$, when conditioned on N_t,

$$x_t(N_t)-D_t(p_t(N_t),N_t)=\Delta_t(N_t)-\xi_t\leq 0<x_{t-1}(N_{t-1})\text{ with probability 1,}$$

i.e., (8) holds for this case. Therefore, for any $\sigma\in[0,\theta]$, the sample-path property (8) holds. This completes the proof of Lemma 1. *Q.E.D.*

Proof of Lemma 4:. By parts (a) and (b) of Theorem 1, if $I_t\leq x_t(N_t)$,

$$v_t(I_t,N_t)=cI_t+\pi_t(N_t),$$

where

$$\pi_t(N_t):=\max\{J_t(x_t,p_t,N_t):x_t\geq 0,p_t\in[\underline{p},\bar{p}]\}.$$

By Lemma 3, $\pi_t(\cdot)$ is concavely increasing and continuously differentiable in N_t.

By Lemma 1, for each $N_t\geq 0$, $x_t(N_t)-D_t(p_t(N_t),N_t)\leq x_{t-1}(N_{t-1})$ with probability 1. Since $v_{t-1}(I_{t-1},N_{t-1})=cI_{t-1}+\pi_{t-1}(N_{t-1})$ for all $I_{t-1}\leq x_{t-1}(N_{t-1})$,

$$\begin{aligned}&v_{t-1}(x_t(N_t)-D_t(p_t(N_t),N_t),\theta D_t(p_t(N_t),N_t)+\eta N_t-\sigma(D_t(p_t(N_t),N_t)-x_t(N_t))^++\epsilon_t)\\ =&c[x_t(N_t)-D_t(p_t(N_t),N_t)]+\pi_{t-1}(\theta D_t(p_t(N_t),N_t)+\eta N_t-\sigma(D_t(p_t(N_t),N_t)-x_t(N_t))^+)\end{aligned}$$

with probability 1. Taking expectation with respect to ϵ_t, we have, for all $N_t\geq 0$ and $x_t\leq x_t(N_t)$,

$$\begin{aligned}&\Psi_t(\Delta_t(N_t)-\xi_t,\theta(y_t(p_t(N_t),N_t)+\xi_t)+\eta N_t-\sigma(\xi_t-\Delta_t(N_t))^+)\\ =&\alpha\mathbb{E}[\pi_{t-1}(\theta(y_t(p_t(N_t),N_t)+\xi_t)+\eta N_t-\sigma(\xi_t-\Delta_t(N_t))^+)].\end{aligned}$$

Therefore, for all $N_t \geq 0$, if $\Delta_t \leq \Delta_t(N_t)$,

$$O_t(\Delta_t, p_t, N_t) = Q_t(p_t, N_t) + \beta\Delta_t + \Lambda(\Delta_t) + \mathbb{E}[G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)],$$

where $G_t(y) := \alpha\mathbb{E}[\pi_{t-1}(y + \epsilon_t)]$.

Finally, it remains to show that $(\Delta_t(N_t), p_t(N_t))$ maximizes the right-hand side of (13). Note that Theorem 1(c) and Lemma 1 imply that, if $I_t \leq x_t(N_t)$, with probability 1, $I_\tau \leq x_\tau(N_\tau)$ for all $\tau = t, t-1, \cdots, 1$ and, hence, $\{(x_\tau(N_\tau), p_\tau(N_\tau))\}_{\tau=t,t-1,\cdots,1}$ is the optimal policy in periods $t, t-1, \cdots, 1$. In particular, $(x_t(N_t), p_t(N_t))$ maximizes the total expected discounted profit given that the firm adopts $\{(x_\tau(N_\tau), p_\tau(N_\tau))\}$ for $\tau = t-1, \cdots, 1$. It's straightforward to check that if the firm adopts the policy $\{(x_\tau(N_\tau), p_\tau(N_\tau))\}$ for $\tau = t-1, \cdots, 1$; and sets the safety-stock level Δ_t and charges p_t in period t, the (normalized) total expected discounted profit of the firm in period t is given by the right-hand side of (13). Since $(\Delta_t(N_t), p_t(N_t))$ maximizes the (normalized) total expected discounted profit in period t, it also maximizes the right-hand side of (13) for each t. This proves Lemma 4. *Q.E.D.*

Proof of Theorem 2: By Theorem 1(c) and Lemma 1, if $I_T \leq x_T(N_t)$, $I_t \leq x_t(N_t)$ for all $t = T, T-1, \cdots, 1$ with probability 1. Therefore, by Theorem 1(a), $(x_t^*(I_t, N_t), p_t^*(I_t, N_t)) = (x_t(N_t), p_t(N_t))$ with probability 1 if $I_T \leq x_T(N_T)$. The characterization of $(\Delta_t(N_t), p_t(N_t))$ follows immediately from Lemma 4 and its discussions. *Q.E.D.*

Before giving the proof of Theorem 3, we first show Theorem 4.

Proof of Theorem 4: Part (a). If $\sigma = 0$, $O_t(\Delta_t, p_t, N_t) = f_1(\Delta_t) + f_2(p_t, N_t)$, where

$$f_1(\Delta_t) := \beta\Delta_t + \Lambda(\Delta_t) \text{ and}$$

$$f_2(p_t, N_t) := Q_t(p_t, N_t) + \mathbb{E}[G_t(\eta N_t + \theta(\bar{V}_t - p_t + \gamma(N_t) + \xi_t))].$$

Since $G_t(\cdot)$ is concave, $f_2(\cdot,\cdot)$ is supermodular in (p_t, N_t). Thus, $p_t(\hat{N}_t) \geq p_t(N_t)$ follows immediately (see Topkis 1998). Now we only consider the case $\sigma > 0$.

Assume, to the contrary, $p_t(\hat{N}_t) < p_t(N_t)$, Lemma 2 implies that $\partial_{p_t} O_t(\Delta_t(\hat{N}_t), p_t(\hat{N}_t), \hat{N}_t) \leq \partial_{p_t} O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}
&\partial_{p_t} Q_t(p_t(\hat{N}_t), \hat{N}_t) - \theta\mathbb{E}[G_t'(\eta\hat{N}_t + \theta(y_t(p_t(\hat{N}_t), \hat{N}_t) + \xi_t) - \sigma(\xi_t - \Delta_t(\hat{N}_t))^+)] \\
\leq &\partial_{p_t} Q_t(p_t(N_t), N_t) - \theta\mathbb{E}[G_t'(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)].
\end{aligned}$$

Since $Q_t(\cdot,\cdot)$ is supermodular in (p_t, N_t) and strictly concave in p_t, $\partial_{p_t} Q_t(p_t(\hat{N}_t), \hat{N}_t) > \partial_{p_t} Q_t(p_t(N_t), N_t)$. Hence,

$$\mathbb{E}[G_t'(\eta\hat{N}_t + \theta(y_t(p_t(\hat{N}_t), \hat{N}_t) + \xi_t) - \sigma(\xi_t - \Delta_t(\hat{N}_t))^+)] > \mathbb{E}[G_t'(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)]. \tag{18}$$

Since $\hat{N}_t > N_t$ and $p_t(\hat{N}_t) < p_t(N_t)$, $y_t(p_t(\hat{N}_t), \hat{N}_t) > y_t(p_t(N_t), N_t)$. The concavity of $G_t(\cdot)$ and (18) imply that $\Delta_t(\hat{N}_t) < \Delta_t(N_t)$. Thus, invoking Lemma 2, we have $\partial_{\Delta_t} O_t(\Delta_t(\hat{N}_t), p_t(\hat{N}_t), \hat{N}_t) \leq \partial_{\Delta_t} O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}
&\beta + \Lambda'(\Delta_t(\hat{N}_t)) + \sigma\mathbb{E}[G_t'(\eta\hat{N}_t + \theta(y_t(p_t(\hat{N}_t), \hat{N}_t) + \xi_t) - \sigma(\xi_t - \Delta_t(\hat{N}_t))^+)1_{\{\xi_t \geq \Delta_t(\hat{N}_t)\}}] \\
\leq &\beta + \Lambda'(\Delta_t(N_t)) + \sigma\mathbb{E}[G_t'(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)1_{\{\xi_t \geq \Delta_t(N_t)\}}].
\end{aligned}$$

The concavity of $\Lambda(\cdot)$ suggests that $\Lambda'(\Delta_t(\hat{N}_t)) \geq \Lambda'(\Delta_t(N_t))$ and, thus,

$$\begin{aligned}&\mathbb{E}[G_t'(\eta\hat{N}_t+\theta(y_t(p_t(\hat{N}_t),\hat{N}_t)+\xi_t)-\sigma(\xi_t-\Delta_t(\hat{N}_t))^+)1_{\{\xi_t\geq\Delta_t(\hat{N}_t)\}}]\\ \leq&\mathbb{E}[G_t'(\eta N_t+\theta(y_t(p_t(N_t),N_t)+\xi_t)-\sigma(\xi_t-\Delta_t(N_t))^+)1_{\{\xi_t\geq\Delta_t(N_t)\}}].\end{aligned} \tag{19}$$

Since $\Delta_t(\hat{N}_t) < \Delta_t(N_t)$, it follows immediately that, for any realization of ξ_t,

$$\begin{aligned}&G_t'(\eta\hat{N}_t+\theta(y_t(p_t(\hat{N}_t),\hat{N}_t)+\xi_t)-\sigma(\xi_t-\Delta_t(\hat{N}_t))^+)1_{\{\xi_t<\Delta_t(\hat{N}_t\}}\\ \leq&G_t'(\eta N_t+\theta(y_t(p_t(N_t),N_t)+\xi_t)-\sigma(\xi_t-\Delta_t(N_t))^+)1_{\{\xi_t<\Delta_t(N_t)\}}.\end{aligned}$$

Integrate over ξ_t and we have

$$\begin{aligned}&\mathbb{E}[G_t'(\eta\hat{N}_t+\theta(y_t(p_t(\hat{N}_t),\hat{N}_t)+\xi_t)-\sigma(\xi_t-\Delta_t(\hat{N}_t))^+)1_{\{\xi_t<\Delta_t(\hat{N}_t)\}}]\\ \leq&\mathbb{E}[G_t'(\eta N_t+\theta(y_t(p_t(N_t),N_t)+\xi_t)-\sigma(\xi_t-\Delta_t(N_t))^+)1_{\{\xi_t<\Delta_t(N_t)\}}].\end{aligned} \tag{20}$$

Sum up (19) and (20) and we have:

$$\mathbb{E}[G_t'(\eta\hat{N}_t+\theta(y_t(p_t(\hat{N}_t),\hat{N}_t)+\xi_t)-\sigma(\xi_t-\Delta_t(\hat{N}_t))^+)]\leq\mathbb{E}[G_t'(\eta N_t+\theta(y_t(p_t(N_t),N_t)+\xi_t)-\sigma(\xi_t-\Delta_t(N_t))^+)],$$

which contradicts (18). Therefore, $p_t(\hat{N}_t) \geq p_t(N_t)$ for all $\hat{N}_t > N_t$. This proves part (a).

Part (b). If $\sigma = 0$, $O_t(\Delta_t, p_t, N_t) = f_1(\Delta_t) + f_2(p_t, N_t)$, so $\Delta_t(\hat{N}_t) = \Delta_t(N_t)$. Now we restrict ourselves to the case $\sigma > 0$.

Assume, to the contrary, that $\Delta_t(\hat{N}_t) > \Delta_t(N_t)$. Lemma 2 implies that $\partial_{\Delta_t} O_t(\Delta_t(\hat{N}_t), p_t(\hat{N}_t), \hat{N}_t) \geq \partial_{\Delta_t} O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}&\beta+\Lambda'(\Delta_t(\hat{N}_t))+\sigma\mathbb{E}[G_t'(\eta\hat{N}_t+\theta(y_t(p_t(\hat{N}_t),\hat{N}_t)+\xi_t)-\sigma(\xi_t-\Delta_t(\hat{N}_t))^+)1_{\{\xi_t\geq\Delta_t(\hat{N}_t)\}}]\\ \geq&\beta+\Lambda'(\Delta_t(N_t))+\sigma\mathbb{E}[G_t'(\eta N_t+\theta(y_t(p_t(N_t),N_t)+\xi_t)-\sigma(\xi_t-\Delta_t(N_t))^+)1_{\{\xi_t\geq\Delta_t(N_t)\}}].\end{aligned}$$

The concavity of $\Lambda(\cdot)$ suggests that $\Lambda'(\Delta_t(\hat{N}_t)) \leq \Lambda'(\Delta_t(N_t))$ and, thus,

$$\begin{aligned}&\mathbb{E}[G_t'(\eta\hat{N}_t+\theta(y_t(p_t(\hat{N}_t),\hat{N}_t)+\xi_t)-\sigma(\xi_t-\Delta_t(\hat{N}_t))^+)1_{\{\xi_t\geq\Delta_t(\hat{N}_t)\}}]\\ \geq&\mathbb{E}[G_t'(\eta N_t+\theta(y_t(p_t(N_t),N_t)+\xi_t)-\sigma(\xi_t-\Delta_t(N_t))^+)1_{\{\xi_t\geq\Delta_t(N_t)\}}].\end{aligned} \tag{21}$$

The concavity of $G_t(\cdot)$ implies that $\eta\hat{N}_t + \theta y_t(p_t(\hat{N}_t), \hat{N}_t) < \eta N_t + \theta y_t(p_t(N_t), N_t)$ and, thus, $p_t(\hat{N}_t) > p_t(N_t)$. Since $\Delta_t(\hat{N}_t) > \Delta_t(N_t)$, it follows immediately that, for any realization of ξ_t,

$$\begin{aligned}&G_t'(\eta\hat{N}_t+\theta(y_t(p_t(\hat{N}_t),\hat{N}_t)+\xi_t)-\sigma(\xi_t-\Delta_t(\hat{N}_t))^+)1_{\{\xi_t<\Delta_t(\hat{N}_t)\}}\\ \geq&G_t'(\eta N_t+\theta(y_t(p_t(N_t),N_t)+\xi_t)-\sigma(\xi_t-\Delta_t(N_t))^+)1_{\{\xi_t<\Delta_t(N_t)\}}.\end{aligned}$$

Integrate over ξ_t and we have

$$\begin{aligned}&\mathbb{E}[G_t'(\eta\hat{N}_t+\theta(y_t(p_t(\hat{N}_t),\hat{N}_t)+\xi_t)-\sigma(\xi_t-\Delta_t(\hat{N}_t))^+)1_{\{\xi_t<\Delta_t(\hat{N}_t)\}}]\\ \geq&\mathbb{E}[G_t'(\eta N_t+\theta(y_t(p_t(N_t),N_t)+\xi_t)-\sigma(\xi_t-\Delta_t(N_t))^+)1_{\{\xi_t<\Delta_t(N_t)\}}].\end{aligned} \tag{22}$$

Sum up (21) and (22) and we have:

$$\mathbb{E}[G_t'(\eta\hat{N}_t+\theta(y_t(p_t(\hat{N}_t),\hat{N}_t)+\xi_t)-\sigma(\xi_t-\Delta_t(\hat{N}_t))^+)]\geq\mathbb{E}[G_t'(\eta N_t+\theta(y_t(p_t(N_t),N_t)+\xi_t)-\sigma(\xi_t-\Delta_t(N_t))^+)]. \tag{23}$$

Since $p_t(\hat{N}_t) > p_t(N_t)$, Lemma 2 implies that $\partial_{p_t} O_t(\Delta_t(\hat{N}_t), p_t(\hat{N}_t), \hat{N}_t) \geq \partial_{p_t} O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}&\partial_{p_t}Q_t(p_t(\hat{N}_t),\hat{N}_t)-\theta\mathbb{E}[G_t'(\eta\hat{N}_t+\theta(y_t(p_t(\hat{N}_t),\hat{N}_t)+\xi_t)-\sigma(\xi_t-\Delta_t(\hat{N}_t))^+)]\\ \geq&\partial_{p_t}Q_t(p_t(N_t),N_t)-\theta\mathbb{E}[G_t'(\eta N_t+\theta(y_t(p_t(N_t),N_t)+\xi_t)-\sigma(\xi_t-\Delta_t(N_t))^+)].\end{aligned}$$

Since $y_t(p_t(\hat{N}_t),\hat{N}_t)<y_t(p_t(N_t),N_t)$,

$$\partial_{p_t}Q_t(p_t(\hat{N}_t),\hat{N}_t)=y_t(p_t(\hat{N}_t),\hat{N}_t)-p_t(\hat{N}_t)+c<y_t(p_t(N_t),N_t)-p_t(N_t)+c=\partial_{p_t}Q_t(p_t(N_t),N_t).$$

Thus,

$$\mathbb{E}[G_t'(\eta\hat{N}_t+\theta(y_t(p_t(\hat{N}_t),\hat{N}_t)+\xi_t)-\sigma(\xi_t-\Delta_t(\hat{N}_t))^+)]<\mathbb{E}[G_t'(\eta N_t+\theta(y_t(p_t(N_t),N_t)+\xi_t)-\sigma(\xi_t-\Delta_t(N_t))^+)],$$

which contradicts inequality (23). Therefore, $\Delta_t(\hat{N}_t)\le\Delta_t(N_t)$ for any $\hat{N}_t>N_t$. This proves part (b).

Part (c). Since $\gamma(\hat{N}_t)=\gamma(N_t)$, $p_t(\hat{N}_t)\ge p_t(N_t)$ implies that $y_t(p_t(\hat{N}_t),\hat{N}_t)\le y_t(p_t(N_t),N_t)$. Moreover, by part (b), $\Delta_t(\hat{N}_t)\le\Delta_t(N_t)$. Therefore,

$$x_t(\hat{N}_t)=\Delta_t(\hat{N}_t)+y_t(p_t(\hat{N}_t),\hat{N}_t)\le\Delta_t(N_t)+y_t(p_t(N_t),N_t)=x_t(N_t).$$

This proves part (c). *Q.E.D.*

Proof of Theorem 3: Part (a). Since $\gamma(\cdot)\equiv\gamma_0$, $\partial_{N_t}v_t(\cdot,\cdot)\equiv 0$, $\pi_t'(\cdot)\equiv 0$, and thus $G_t'(\cdot)\equiv 0$. Therefore, optimizing (9) yields that $\Delta_t(N_t)\equiv\Delta_*$ for any t and N_t. To show $\hat{\Delta}_t(N_t)\ge\Delta_*$, we assume, to the contrary, that $\hat{\Delta}_t(N_t)<\Delta_*$. Lemma 2 yields that $\partial_{\Delta_t}\hat{O}_t(\hat{\Delta}_t(N_t),\hat{p}_t(N_t),N_t)\le\beta+\Lambda'(\Delta_*)$, i.e.,

$$\beta+\Lambda'(\hat{\Delta}_t(N_t))+\sigma\mathbb{E}[G_t'(\eta N_t+\theta(y_t(\hat{p}_t(N_t),N_t)+\xi_t)-\sigma(\xi_t-\hat{\Delta}_t(N_t))^+)1_{\{\xi_t\ge\hat{\Delta}_t(N_t)\}}]\le\beta+\Lambda'(\Delta_*).$$

Since $\sigma\mathbb{E}[G_t'(\eta N_t+\theta(y_t(\hat{p}_t(N_t),N_t)+\xi_t)-\sigma(\xi_t-\hat{\Delta}_t(N_t))^+)1_{\{\xi_t\ge\hat{\Delta}_t(N_t)\}}]\ge 0$, we have $\Lambda'(\hat{\Delta}_t(N_t))\le\Lambda'(\Delta_*)$. The concavity of $\Lambda(\cdot)$ indicates that $\Lambda'(\hat{\Delta}_t(N_t))=\Lambda'(\Delta_*)$ and thus, by our assumption that Δ_* is the lexicographically smallest optimizer, $\hat{\Delta}_t(N_t)\ge\Delta_*$. This contradicts with $\hat{\Delta}_t(N_t)<\Delta_*$. Thus, $\hat{\Delta}_t(N_t)\ge\Delta_*$.

If $\hat{\gamma}'(\cdot)>0$, $\hat{G}_t'(\cdot)>0$. Thus, for any p_t,

$$\begin{aligned}\partial_{\Delta_t}\hat{O}_t(\Delta_*,p_t,N_t)=&\beta+\Lambda'(\Delta_*)+\sigma\mathbb{E}[\hat{G}_t'(\eta N_t+\theta(y_t(p_t,N_t)+\xi_t)-\sigma(\xi_t-\Delta_*)^+)1_{\{\xi_t\ge\hat{\Delta}_t(N_t)\}}]\\=&\sigma\mathbb{E}[G_t'(\eta N_t+\theta(y_t(p_t,N_t)+\xi_t)-\sigma(\xi_t-\Delta_*)^+)1_{\{\xi_t\ge\hat{\Delta}_t(N_t)\}}]\\>&0,\end{aligned}$$

where the second equality follows from the first-order condition $\beta+\Lambda'(\Delta_*)=0$ and the inequality from $\sigma>0$ and $\hat{G}'(\cdot)>0$. Hence, $\hat{\Delta}_t(N_t)>\Delta_*$. This proves part (a).

Part (b). We rewrite the objective function (9) in (Δ_t,y_t,N_t), where $y_t=y_t(p_t,N_t)=\bar{V}_t-p_t+\gamma(N_t)$ is the expected demand in period t. It is clear that, given the net rating N_t, price p_t and expected demand y_t have a one-to-one correspondence. Hence, optimizing over (Δ_t,p_t,N_t) is equivalent to optimizing over (Δ_t,y_t,N_t). We transform the objective function $O_t(\cdot,\cdot,\cdot)$ into

$$K_t(\Delta_t,y_t,N_t)=(\bar{V}_t+\gamma(N_t)-y_t-c)y_t+\beta\Delta_t+\Lambda(\Delta_t)+\mathbb{E}[G_t(\eta N_t+\theta(y_t+\xi_t)-\sigma(\xi_t-\Delta_t)^+)].$$

Let $y_t(N_t)$ be the optimal expected demand in period t with net rating N_t. We have $y_t(N_t)=y_t(p_t(N_t),N_t)=\bar{V}_t-p_t(N_t)+\gamma(N_t)$.

We now show that $\hat{y}_t(N_t)\ge y_t(N_t)$ for all N_t. Assume, to the contrary, that $\hat{y}_t(N_t)<y_t(N_t)$. Lemma 2 yields that $\partial_{y_t}\hat{K}_t(\hat{\Delta}_t(N_t),\hat{y}_t(N_t),N_t)\le\partial_{y_t}K_t(\Delta_t(N_t),y_t(N_t),N_t)$, i.e.,

$$\begin{aligned}&\bar{V}_t-c-2\hat{y}_t(N_t)+\hat{\gamma}(N_t)+\theta\mathbb{E}[\hat{G}_t'(\eta N_t+\theta(\hat{y}_t(N_t)+\xi_t)-\sigma(\xi_t-\hat{\Delta}_t(N_t))^+)]\\\le&\bar{V}_t-c-2y_t(N_t)+\gamma(N_t)+\theta\mathbb{E}[G_t'(\eta N_t+\theta(y_t(N_t)+\xi_t)-\sigma(\xi_t-\Delta_t(N_t))^+)].\end{aligned}$$

Because $\hat{G}'_t(\cdot) \geq G'_t(\cdot) \equiv 0$ and $\hat{y}_t(N_t) < y_t(N_t)$, the concavity of $\hat{G}_t(\cdot)$ implies that

$$\mathbb{E}[\hat{G}'_t(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)] \geq \mathbb{E}[G'_t(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)].$$

Since $\hat{\gamma}(N_t) \geq \gamma(N_t) \equiv 0$, we have $-2\hat{y}_t(N_t) \leq -2y_t(N_t)$, which contradicts the assumption that $\hat{y}_t(N_t) < y_t(N_t)$. Hence, $\hat{y}_t(N_t) \geq y_t(N_t)$.

We now show $y_t(N_t) \equiv y_*(t)$. Observe that, since $\gamma(\cdot) \equiv 0$ and $G'_t(\cdot) \equiv 0$, optimizing (9) yields that

$$p_t(N_t) = \arg\max_{p_t \in [\underline{p}, \bar{p}]} Q_t(p_t, N_t) = \min\{\max\{\frac{\bar{V}_t + c}{2}, \underline{p}\}, \bar{p}\} =: p_*(t).$$

Clearly, $p_*(t)$ is independent of the net rating N_t. Hence, $y_t(N_t) = \bar{V}_t - p_t(N_t) + \gamma(N_t) = \bar{V}_t - p_*(t) =: y_*(t)$, which is independent of the net rating N_t as well.

Putting everything together, we have

$$\hat{x}_t(N_t) = \hat{y}_t(N_t) + \hat{\Delta}_t(N_t) \geq y_t(N_t) + \Delta_t(N_t) = y_*(t) + \Delta_* =: x_*(t).$$

Here, $x_*(t)$ is independent of the net rating N_t. By part (a), the inequality is strict if $\sigma > 0$ and $\gamma'(\cdot) > 0$. This proves part (b).

Part (c). The equality $p_t(N_t) \equiv p_*(t)$ has been shown in part (b). To show the existence of the threshold $\mathfrak{N}_t$, we first prove that $\hat{p}_t(0) \leq p_t(0)$. Observe that $\hat{p}_t(0) = \bar{V}_t + \hat{\gamma}(0) - \hat{y}_t(0)$ and $p_t(0) = \bar{V}_t + \gamma(0) - y_t(0)$. By the proof of part (b), $\hat{y}_t(0) \geq y_t(0)$. Moreover, since $\hat{\gamma}(0) = \gamma(0) = 0$, $\hat{p}_t(0) \leq p_t(0)$. By part (b), $p_t(N_t) \equiv p_t(0)$ for all N_t. Recall from Theorem 4(a) that $\hat{p}_t(N_t)$ is increasing in N_t. The joint concavity of $\hat{O}_t(\cdot,\cdot,\cdot)$ implies that $\hat{p}_t(N_t)$ is continuously increasing in N_t. Thus, let $\mathfrak{N}_t$ be the smallest N_t such that $\hat{p}_t(N_t) \geq p_t(N_t) = p_*(t)$. The monotonicity of $\hat{p}_t(\cdot)$ then suggests that $\hat{p}_t(N_t) \leq p_t(N_t) \equiv p_*(t)$ if $N_t \leq \mathfrak{N}_t$, and $\hat{p}_t(N_t) \geq p_t(N_t) \equiv p_*(t)$ if $N_t \geq \mathfrak{N}_t$. This proves part (c). *Q.E.D.*

Proof of Theorem 5: We show Theorem 5 by backward induction. More specifically, we show that if $\bar{V}_\tau \equiv \bar{V}$ for all τ and $\pi'_{t-1}(N) \geq \pi'_{t-2}(N)$ for all $N \geq 0$, (i) $p_t(N) \leq p_{t-1}(N)$ for all $N \geq 0$, (ii) $\Delta_t(N) \geq \Delta_{t-1}(N)$ for all $N \geq 0$, (iii) $x_t(N) \geq x_{t-1}(N)$ for all $N \geq 0$, and (iv) $\pi'_t(N) \geq \pi'_{t-1}(N)$ for all $N \geq 0$. Since $\pi'_1(N) \geq \pi'_0(N) \equiv 0$ for all N, the initial condition is satisfied.

Note that $\pi'_{t-1}(N) \geq \pi'_{t-2}(N)$ for all $N \geq 0$ implies that

$$G'_t(y) = \alpha \mathbb{E}[\pi'_{t-1}(y + \epsilon_t)] \geq \alpha \mathbb{E}[\pi'_{t-2}(y + \epsilon_t)] = G'_{t-1}(y),$$

for all y. Since $\bar{V}_t \equiv \bar{V}$, $Q_t(p_t, N_t) = (p_t - c)(\bar{V} - p_t + \gamma(N_t)) =: Q(p_t, N_t)$ for all t. We use $y_t(N) := y_t(p_t(N), N)$ to denote the expected demand in period t with net rating N under the optimal policy.

We first prove that $p_t(N) \leq p_{t-1}(N)$ for all N. Assume, to the contrary, that $p_t(N) > p_{t-1}(N)$ for some N. Lemma 2 implies that $\partial_{p_t} O_t(\Delta_t(N), p_t(N), N) \geq \partial_{p_{t-1}} O_{t-1}(\Delta_{t-1}(N), p_{t-1}(N), N)$, i.e.,

$$\begin{aligned}&\partial_p Q(p_t(N), N) - \theta \mathbb{E}[G'_t(\eta N + \theta(y_t(N) + \xi_t) - \sigma(\xi_t - \Delta_t(N))^+)]\\ \geq&\partial_p Q(p_{t-1}(N), N) - \theta \mathbb{E}[G'_{t-1}(\eta N + \theta(y_{t-1}(N) + \xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)].\end{aligned}$$

Since $Q(\cdot, N)$ is strictly concave in p and $p_t(N) > p_{t-1}(N)$, $\partial_p Q(p_t(N), N) < \partial_p Q(p_{t-1}(N), N)$. Thus,

$$\mathbb{E}[G'_t(\eta N + \theta(y_t(N) + \xi_t) - \sigma(\xi_t - \Delta_t(N))^+)] < \mathbb{E}[G'_{t-1}(\eta N + \theta(y_{t-1}(N) + \xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)]. \tag{24}$$

Note that $G'_t(\cdot) \geq G'_{t-1}(\cdot)$ for all y, $p_t(N) > p_{t-1}(N)$, and $G_t(\cdot)$ and $G_{t-1}(\cdot)$ are concave. We have (24) implies that $\sigma > 0$ and $\Delta_t(N) > \Delta_{t-1}(N)$. Thus, Lemma 2 implies that $\partial_{\Delta_t} O_t(\Delta_t(N), p_t(N), N) \geq \partial_{\Delta_{t-1}} O_{t-1}(\Delta_{t-1}(N), p_{t-1}(N), N)$, i.e.,

$$\begin{aligned}&\beta + \Lambda'(\Delta_t(N)) + \sigma\mathbb{E}[G'_t(\eta N + \theta(y_t(N)+\xi_t) - \sigma(\xi_t - \Delta_t(N))^+)1_{\{\xi_t \geq \Delta_t(N)\}}]\\ \geq&\beta + \Lambda'(\Delta_{t-1}(N)) + \sigma\mathbb{E}[G'_{t-1}(\eta N + \theta(y_{t-1}(N)+\xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)1_{\{\xi_{t-1} \geq \Delta_{t-1}(N)\}}].\end{aligned}$$

The concavity of $\Lambda(\cdot)$ suggests that $\Lambda'(\Delta_t(N_t)) \leq \Lambda'(\Delta_{t-1}(N))$ and, thus,

$$\begin{aligned}&\mathbb{E}[G'_t(\eta N + \theta(y_t(N)+\xi_t) - \sigma(\xi_t - \Delta_t(N))^+)1_{\{\xi_t \geq \Delta_t(N)\}}]\\ \geq&\mathbb{E}[G'_{t-1}(\eta N + \theta(y_{t-1}(N)+\xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)1_{\{\xi_{t-1} \geq \Delta_{t-1}(N)\}}].\end{aligned} \tag{25}$$

Since $\Delta_t(N) > \Delta_{t-1}(N)$ and $G'_t(\cdot) \geq G'_{t-1}(\cdot) \geq 0$, it follows immediately that, for any realization of $\xi_t = \xi_{t-1}$,

$$\begin{aligned}&G'_t(\eta N + \theta(y_t(N)+\xi_t) - \sigma(\xi_t - \Delta_t(N))^+)1_{\{\xi_t < \Delta_t(N)\}}\\ \geq&G'_{t-1}(\eta N + \theta(y_{t-1}(N)+\xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)1_{\{\xi_{t-1} < \Delta_{t-1}(N)\}}.\end{aligned}$$

Integrate over ξ_t and ξ_{t-1} and we have

$$\begin{aligned}&\mathbb{E}[G'_t(\eta N + \theta(y_t(N)+\xi_t) - \sigma(\xi_t - \Delta_t(N))^+)1_{\{\xi_t < \Delta_t(N)\}}]\\ \geq&\mathbb{E}[G'_{t-1}(\eta N + \theta(y_{t-1}(N)+\xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)1_{\{\xi_{t-1} < \Delta_{t-1}(N)\}}].\end{aligned} \tag{26}$$

Sum up (25) and (26) and we have:

$$\mathbb{E}[G'_t(\eta N + \theta(y_t(N)+\xi_t) - \sigma(\xi_t - \Delta_t(N))^+)] \geq \mathbb{E}[G'_{t-1}(\eta N + \theta(y_{t-1}(N)+\xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)],$$

which contradicts (24). Therefore, $p_t(N) \leq p_{t-1}(N)$ for all N.

Next, we show that $\Delta_t(N) \geq \Delta_{t-1}(N)$. If $\sigma = 0$, it is straightforward to show that $\Delta_t(N) = \Delta_{t-1}(N) = \Delta_*$. Hence, we confine ourselves to the interesting case of $\sigma > 0$.

Assume, to the contrary, that $\Delta_t(N) < \Delta_{t-1}(N)$. Lemma 2 implies that $\partial_{\Delta_t} O_t(\Delta_t(N), p_t(N), N) \leq \partial_{\Delta_{t-1}} O_{t-1}(\Delta_{t-1}(N), p_{t-1}(N), N)$, i.e.,

$$\begin{aligned}&\beta + \Lambda'(\Delta_t(N)) + \sigma\mathbb{E}[G'_t(\eta N + \theta(y_t(N)+\xi_t) - \sigma(\xi_t - \Delta_t(N))^+)1_{\{\xi_t \geq \Delta_t(N)\}}]\\ \leq&\beta + \Lambda'(\Delta_{t-1}(N)) + \sigma\mathbb{E}[G'_{t-1}(\eta N + \theta(y_{t-1}(N)+\xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)1_{\{\xi_{t-1} \geq \Delta_{t-1}(N)\}}].\end{aligned}$$

The concavity of $\Lambda(\cdot)$ suggests that $\Lambda'(\Delta_t(N)) \geq \Lambda'(\Delta_{t-1}(N))$ and, thus,

$$\begin{aligned}&\mathbb{E}[G'_t(\eta N + \theta(y_t(N)+\xi_t) - \sigma(\xi_t - \Delta_t(N))^+)1_{\{\xi_t \geq \Delta_t(N)\}}]\\ \leq&\mathbb{E}[G'_{t-1}(\eta N + \theta(y_{t-1}(N)+\xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)1_{\{\xi_{t-1} \geq \Delta_{t-1}(N)\}}].\end{aligned} \tag{27}$$

The concavity of $G_t(\cdot)$ and $G_{t-1}(\cdot)$ and that $G'_t(\cdot) \geq G'_{t-1}(\cdot)$ imply that $y_t(N) > y_{t-1}(N_t)$ and, thus, $p_t(N) < p_{t-1}(N)$. Since $\Delta_t(N) < \Delta_{t-1}(N)$, it follows immediately that, for any realization of $\xi_t = \xi_{t-1}$,

$$\begin{aligned}&G'_t(\eta N + \theta(y_t(N)+\xi_t) - \sigma(\xi_t - \Delta_t(N))^+)1_{\{\xi_t < \Delta_t(N)\}}\\ \leq&G'_{t-1}(\eta N + \theta(y_{t-1}(N)+\xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)1_{\{\xi_{t-1} < \Delta_{t-1}(N)\}}.\end{aligned}$$

Integrate over ξ_t and ξ_{t-1} and we have

$$\begin{aligned}&\mathbb{E}[G'_t(\eta N + \theta(y_t(N)+\xi_t) - \sigma(\xi_t - \Delta_t(N))^+)1_{\{\xi_t < \Delta_t(N)\}}]\\ \leq&\mathbb{E}[G'_{t-1}(\eta N + \theta(y_{t-1}(N)+\xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)1_{\{\xi_{t-1} < \Delta_{t-1}(N)\}}].\end{aligned} \tag{28}$$

Sum up (27) and (28) and we have:

$$\mathbb{E}[G_t'(\eta N+\theta(y_t(N)+\xi_t)-\sigma(\xi_t-\Delta_t(N))^+)]\leq\mathbb{E}[G_{t-1}'(\eta N+\theta(y_{t-1}(N)+\xi_{t-1})-\sigma(\xi_{t-1}-\Delta_{t-1}(N))^+)]. \tag{29}$$

By Lemma 2, $p_t(N)<p_{t-1}(N)$ yields that $\partial_{p_t}O_t(\Delta_t(N),p_t(N),N)\leq\partial_{p_{t-1}}O_{t-1}(\Delta_{t-1}(N),p_{t-1}(N),N)$, i.e.,

$$\begin{aligned}&\partial_pQ(p_t(N),N)-\theta\mathbb{E}[G_t'(\eta N+\theta(y_t(N)+\xi_t)-\sigma(\xi_t-\Delta_t(N))^+)]\\ \leq&\partial_pQ(p_{t-1}(N),N)-\theta\mathbb{E}[G_{t-1}'(\eta N+\theta(y_{t-1}(N)+\xi_{t-1})-\sigma(\xi_{t-1}-\Delta_{t-1}(N))^+)].\end{aligned}$$

Since $Q(\cdot,N)$ is strictly concave in p, $\partial_pQ(p_t(N),N)>\partial_pQ(p_{t-1}(N),N)$. Thus,

$$\mathbb{E}[G_t'(\eta N+\theta(y_t(N)+\xi_t)-\sigma(\xi_t-\Delta_t(N))^+)]>\mathbb{E}[G_{t-1}'(\eta N+\theta(y_{t-1}(N)+\xi_{t-1})-\sigma(\xi_{t-1}-\Delta_{t-1}(N))^+)],$$

which contradicts inequality (29). Therefore, $\Delta_t(N)\geq\Delta_{t-1}(N)$ for any N.

Next, we show $x_t(N)\geq x_{t-1}(N)$. Note that $p_t(N)\leq p_{t-1}(N)$ implies that $y_t(N)\geq y_{t-1}(N)$. Thus,

$$x_t(N)=y_t(N)+\Delta_t(N)\geq y_{t-1}(N)+\Delta_{t-1}(N)=x_{t-1}(N).$$

Finally, to complete the induction, we show that $\pi_t'(N)\geq\pi_{t-1}'(N)$ for all N. By the envelope theorem,

$$\pi_t'(N)=(p_t(N)-c)\gamma'(N)+(\eta+\theta\gamma'(N))\mathbb{E}[G_t'(\eta N+\theta(y_t(N)+\xi_t)-\sigma(\xi_t-\Delta_t(N))^+)],$$

and

$$\pi_{t-1}'(N)=(p_{t-1}(N)-c)\gamma'(N)+(\eta+\theta\gamma'(N))\mathbb{E}[G_{t-1}'(\eta N+\theta(y_{t-1}(N)+\xi_{t-1})-\sigma(\xi_{t-1}-\Delta_{t-1}(N))^+)],$$

If $p_t(N)=p_{t-1}(N)$ and $\Delta_t(N)=\Delta_{t-1}(N)$, $\pi_t'(N)\geq\pi_{t-1}'(N)$ follows immediately from $\gamma'(N)\geq0$ and $G_t'(\cdot)\geq G_{t-1}'(\cdot)$.

If $p_t(N)=p_{t-1}(N)$ and $\Delta_t(N)>\Delta_{t-1}(N)$, Lemma 2 yields that $\partial_{\Delta_t}O_t(\Delta_t(N),p_t(N),N)\geq\partial_{\Delta_{t-1}}O_{t-1}(\Delta_{t-1}(N),p_{t-1}(N),N)$, i.e.,

$$\begin{aligned}&\beta+\Lambda'(\Delta_t(N))+\sigma\mathbb{E}[G_t'(\eta N+\theta(y_t(N)+\xi_t)-\sigma(\xi_t-\Delta_t(N))^+)1_{\{\xi_t\geq\Delta_t(N)\}}]\\ \geq&\beta+\Lambda'(\Delta_{t-1}(N))+\sigma\mathbb{E}[G_{t-1}'(\eta N+\theta(y_{t-1}(N)+\xi_{t-1})-\sigma(\xi_{t-1}-\Delta_{t-1}(N))^+)1_{\{\xi_{t-1}\geq\Delta_{t-1}(N)\}}].\end{aligned}$$

The concavity of $\Lambda(\cdot)$ suggests that $\Lambda'(\Delta_t(N))\leq\Lambda'(\Delta_{t-1}(N))$ and, thus,

$$\begin{aligned}&\mathbb{E}[G_t'(\eta N+\theta(y_t(N)+\xi_t)-\sigma(\xi_t-\Delta_t(N))^+)1_{\{\xi_t\geq\Delta_t(N)\}}]\\ \geq&\mathbb{E}[G_{t-1}'(\eta N+\theta(y_{t-1}(N)+\xi_{t-1})-\sigma(\xi_{t-1}-\Delta_{t-1}(N))^+)1_{\{\xi_{t-1}\geq\Delta_{t-1}(N)\}}].\end{aligned} \tag{30}$$

Since $\Delta_t(N)>\Delta_{t-1}(N)$, it follows immediately that, for any realization of $\xi_t=\xi_{t-1}$,

$$\begin{aligned}&G_t'(\eta N+\theta(y_t(N)+\xi_t)-\sigma(\xi_t-\Delta_t(N))^+)1_{\{\xi_t<\Delta_t(N)\}}\\ \geq&G_{t-1}'(\eta N+\theta(y_{t-1}(N)+\xi_{t-1})-\sigma(\xi_{t-1}-\Delta_{t-1}(N))^+)1_{\{\xi_{t-1}<\Delta_{t-1}(N)\}}.\end{aligned}$$

Integrate over ξ_t and ξ_{t-1} and we have

$$\begin{aligned}&\mathbb{E}[G_t'(\eta N+\theta(y_t(N)+\xi_t)-\sigma(\xi_t-\Delta_t(N))^+)1_{\{\xi_t<\Delta_t(N)\}}]\\ \geq&\mathbb{E}[G_{t-1}'(\eta N+\theta(y_{t-1}(N)+\xi_{t-1})-\sigma(\xi_{t-1}-\Delta_{t-1}(N))^+)1_{\{\xi_{t-1}<\Delta_{t-1}(N)\}}].\end{aligned} \tag{31}$$

Sum up (30) and (31) and we have:

$$\mathbb{E}[G_t'(\eta N+\theta(y_t(N)+\xi_t)-\sigma(\xi_t-\Delta_t(N))^+)]\geq\mathbb{E}[G_{t-1}'(\eta N+\theta(y_{t-1}(N)+\xi_{t-1})-\sigma(\xi_{t-1}-\Delta_{t-1}(N))^+)]. \tag{32}$$

Plugging (32) into the formulas of $\pi'_t(\cdot)$ and $\pi'_{t-1}(\cdot)$, we have that the inequality $\pi'_t(N) \geq \pi'_{t-1}(N)$ follows immediately from $p_t(N) = p_{t-1}(N)$.

If $p_t(N) < p_{t-1}(N)$, Lemma 2 yields that $\partial_{p_t} O_t(\Delta_t(N), p_t(N), N) \leq \partial_{p_{t-1}} O_{t-1}(\Delta_{t-1}(N), p_{t-1}(N), N)$, i.e.,

$$\begin{aligned}&\partial_p Q(p_t(N), N) - \theta\mathbb{E}[G'_t(\eta N + \theta(y_t(N) + \xi_t) - \sigma(\xi_t - \Delta_t(N))^+)]\\ \leq&\partial_p Q(p_{t-1}(N), N) - \theta\mathbb{E}[G'_{t-1}(\eta N + \theta(y_{t-1}(N) + \xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)],\end{aligned}$$

i.e.,

$$\begin{aligned}&\bar{V} + c - 2p_t(N) + \gamma(N) - \theta\mathbb{E}[G'_t(\eta N + \theta(y_t(N) + \xi_t) - \sigma(\xi_t - \Delta_t(N))^+)]\\ \leq&\bar{V} + c - 2p_{t-1}(N) + \gamma(N) - \theta\mathbb{E}[G'_{t-1}(\eta N + \theta(y_{t-1}(N) + \xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)].\end{aligned}$$

Thus,

$$\begin{aligned}&(p_t(N) - p_{t-1}(N)) + \theta(\mathbb{E}[G'_t(\eta N + \theta(y_t(N) + \xi_t) - \sigma(\xi_t - \Delta_t(N))^+)]\\ &- \mathbb{E}[G'_{t-1}(\eta N + \theta(y_{t-1}(N) + \xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)])\\ \geq&p_{t-1}(N) - p_t(N)\\ >&0.\end{aligned}$$

Moreover,

$$\begin{aligned}&\mathbb{E}[G'_t(\eta N + \theta(y_t(N) + \xi_t) - \sigma(\xi_t - \Delta_t(N))^+)]\\ &- \mathbb{E}[G'_{t-1}(\eta N + \theta(y_{t-1}(N) + \xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)]\\ \geq&\frac{2}{\theta}(p_{t-1}(N) - p_t(N))\\ >&0.\end{aligned}$$

Therefore,

$$\begin{aligned}\pi'_t(N) - \pi'_{t-1}(N) =&((p_t(N) - p_{t-1}(N)) + \theta(\mathbb{E}[G'_t(\eta N + \theta(y_t(N) + \xi_t) - \sigma(\xi_t - \Delta_t(N))^+)]\\ &- \mathbb{E}[G'_{t-1}(\eta N + \theta(y_{t-1}(N) + \xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)])\gamma'(N)\\ &+ \eta(\mathbb{E}[G'_t(\eta N + \theta(y_t(N) + \xi_t) - \sigma(\xi_t - \Delta_t(N))^+)]\\ &- \mathbb{E}[G'_{t-1}(\eta N + \theta(y_{t-1}(N) + \xi_{t-1}) - \sigma(\xi_{t-1} - \Delta_{t-1}(N))^+)])\\ \geq&0.\end{aligned}$$

Hence, $\pi'_t(N) \geq \pi'_{t-1}(N)$ for all N. This completes the induction and, thus, the proof of Theorem 5. *Q.E.D.*

Proof of Theorem 6: We show Theorem 6 by backward induction. More specifically, we show that if $\hat{\alpha} > \alpha$ and $\hat{\pi}'_{t-1}(N_{t-1}) \geq \pi'_{t-1}(N_{t-1})$ for all $N_{t-1} \geq 0$, (i) $\hat{p}_t(N_t) \leq p_t(N_t)$ for all $N_t \geq 0$, (ii) $\hat{\Delta}_t(N_t) \geq \Delta_t(N_t)$ for all $N_t \geq 0$, (iii) $\hat{x}_t(N_t) \geq x_t(N_t)$ for all $N_t \geq 0$, and (iv) $\hat{\pi}'_t(N_t) \geq \pi'_t(N_t)$ for all $N_t \geq 0$. Since $\hat{\pi}'_0(N_0) = \pi'_0(N_0) \equiv 0$ for all N, the initial condition is satisfied.

Note that $\hat{\pi}'_{t-1}(N_{t-1}) \geq \pi'_{t-1}(N_{t-1})$ for all $N_{t-1} \geq 0$ implies that

$$\hat{G}'_t(y) = \hat{\alpha}\mathbb{E}[\hat{\pi}'_{t-1}(y + \epsilon_t)] \geq \alpha\mathbb{E}[\pi'_{t-1}(y + \epsilon_t)] = G'_t(y),$$

for all y. We use $\hat{y}_t(N_t) := y_t(\hat{p}_t(N_t), N_t)$ and $y_t(N_t) := y_t(p_t(N_t), N_t)$ to denote the expected demand in period t under the optimal policy with discount factor $\hat{\alpha}$ and α, respectively. Since $\hat{\alpha} > \alpha$, $\hat{\beta} = b - (1 - \hat{\alpha})c > b - (1 - \alpha)c = \beta$.

We first prove that $\hat{p}_t(N_t) \le p_t(N_t)$ for all N_t. Assume, to the contrary, that $\hat{p}_t(N_t) > p_t(N_t)$ for some N_t. Lemma 2 implies that $\partial_{p_t} O_t(\hat{\Delta}_t(N_t), \hat{p}_t(N_t), N_t) \ge \partial_{p_t} O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}&\partial_{p_t} Q_t(\hat{p}_t(N_t), N_t) - \theta \mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)]\\ \ge&\partial_{p_t} Q_t(p_t(N_t), N_t) - \theta \mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)].\end{aligned}$$

Since $Q_t(\cdot, N_t)$ is strictly concave in p_t and $\hat{p}_t(N_t) > p_t(N_t)$, $\partial_{p_t} \hat{Q}_t(p_t(N_t), N_t) < \partial_{p_t} Q(p_t(N_t), N_t)$. Thus,

$$\mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)] < \mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)]. \tag{33}$$

Note that $\hat{G}_t'(\cdot) \ge G_t'(\cdot)$ for all y, $\hat{p}_t(N_t) > p_t(N_t)$, and $\hat{G}_t(\cdot)$ and $G_t(\cdot)$ are concave. We have (33) implies that $\sigma > 0$ and $\hat{\Delta}_t(N_t) > \Delta_t(N_t)$. Thus, Lemma 2 implies that $\partial_{\Delta_t} \hat{O}_t(\hat{\Delta}_t(N_t), \hat{p}_t(N_t), N_t) \ge \partial_{\Delta_t} O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}&\hat{\beta} + \Lambda'(\hat{\Delta}_t(N_t)) + \sigma \mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+) 1_{\{\xi_t \ge \hat{\Delta}_t(N_t)\}}]\\ \ge&\beta + \Lambda'(\Delta_t(N_t)) + \sigma \mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+) 1_{\{\xi_t \ge \Delta_t(N_t)\}}].\end{aligned}$$

The concavity of $\Lambda(\cdot)$ suggests that $\Lambda'(\hat{\Delta}_t(N_t)) \le \Lambda'(\Delta_t(N_t))$. In addition, $\hat{\beta} > \beta$, thus,

$$\begin{aligned}&\mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+) 1_{\{\xi_t \ge \hat{\Delta}_t(N_t)\}}]\\ \ge&\mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+) 1_{\{\xi_t \ge \Delta_t(N_t)\}}].\end{aligned} \tag{34}$$

Since $\hat{\Delta}_t(N_t) > \Delta_t(N_t)$ and $\hat{G}_t'(\cdot) \ge G_t'(\cdot) \ge 0$, it follows immediately that, for any realization of ξ_t,

$$\begin{aligned}&\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+) 1_{\{\xi_t < \hat{\Delta}_t(N_t)\}}\\ \ge&G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+) 1_{\{\xi_t < \Delta_t(N_t)\}}.\end{aligned}$$

Integrate over ξ_t and we have

$$\begin{aligned}&\mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+) 1_{\{\xi_t < \hat{\Delta}_t(N_t)\}}]\\ \ge&\mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+) 1_{\{\xi_t < \Delta_t(N_t)\}}].\end{aligned} \tag{35}$$

Sum up (34) and (35) and we have:

$$\mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)] \ge \mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)],$$

which contradicts (33). Therefore, $\hat{p}_t(N_t) \le p_t(N_t)$ for all N.

Next, we show that $\hat{\Delta}_t(N_t) \ge \Delta_t(N_t)$. If $\sigma = 0$, $\hat{\Delta}_t(N_t) = \arg\max_{\Delta_t}[\hat{\beta}\Delta_t + L(\Delta_t)]$, whereas $\Delta_t(N_t) = \arg\max_{\Delta_t}[\beta\Delta_t + L(\Delta_t)]$. Since $\hat{\beta} > \beta$, it follows immediately that $\hat{\Delta}_t(N_t) \ge \Delta_t(N_t)$. Hence, we confine ourselves to the interesting case of $\sigma > 0$.

Assume, to the contrary, that $\hat{\Delta}_t(N_t) < \Delta_t(N_t)$. Lemma 2 implies that $\partial_{\Delta_t} \hat{O}_t(\hat{\Delta}_t(N_t), \hat{p}_t(N_t), N_t) \le \partial_{\Delta_t} O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}&\hat{\beta} + \Lambda'(\hat{\Delta}_t(N_t)) + \sigma \mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+) 1_{\{\xi_t \ge \hat{\Delta}_t(N_t)\}}]\\ \le&\beta + \Lambda'(\Delta_t(N_t)) + \sigma \mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+) 1_{\{\xi_t \ge \Delta_t(N_t)\}}].\end{aligned}$$

The concavity of $\Lambda(\cdot)$ suggests that $\Lambda'(\hat{\Delta}_t(N_t)) \ge \Lambda'(\Delta_t(N_t))$. Since $\hat{\beta} > \beta$, we have

$$\begin{aligned}&\mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+) 1_{\{\xi_t \ge \hat{\Delta}_t(N_t)\}}]\\ \le&\mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+) 1_{\{\xi_t \ge \Delta_t(N_t)\}}].\end{aligned} \tag{36}$$

The concavity of $\hat{G}_t(\cdot)$ and $G_t(\cdot)$ and that $\hat{G}'_t(\cdot) \geq G'_t(\cdot)$ imply that $\hat{y}_t(N_t) > y_t(N_t)$ and, thus, $\hat{p}_t(N) < p_t(N_t)$. Since $\hat{\Delta}_t(N_t) < \Delta_t(N_t)$, it follows immediately that, for any realization of ξ_t,

$$\begin{aligned}&\hat{G}'_t(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)1_{\{\xi_t < \hat{\Delta}_t(N_t)\}}\\ \leq& G'_t(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)1_{\{\xi_t < \Delta_t(N_t)\}}.\end{aligned}$$

Integrate over ξ_t and we have

$$\begin{aligned}&\mathbb{E}[\hat{G}'_t(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)1_{\{\xi_t < \hat{\Delta}_t(N_t)\}}]\\ \leq& \mathbb{E}[G'_t(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)1_{\{\xi_t < \Delta_t(N_t)\}}].\end{aligned} \tag{37}$$

Sum up (36) and (37) and we have:

$$\begin{aligned}&\mathbb{E}[\hat{G}'_t(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)]\\ \leq& \mathbb{E}[G'_t(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)].\end{aligned} \tag{38}$$

By Lemma 2, $\hat{p}_t(N_t) < p_t(N_t)$ yields that $\partial_{p_t}\hat{O}_t(\hat{\Delta}_t(N_t), \hat{p}_t(N_t), N_t) \leq \partial_{p_t}O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}&\partial_{p_t}\hat{Q}(\hat{p}_t(N_t), N_t) - \theta\mathbb{E}[\hat{G}'_t(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)]\\ \leq& \partial_{p_t}Q(p_t(N_t), N_t) - \theta\mathbb{E}[G'_t(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)].\end{aligned}$$

Since $Q_t(\cdot, N_t)$ is strictly concave in p_t, $\partial_{p_t}Q_t(\hat{p}_t(N_t), N_t) > \partial_{p_t}Q_t(p_t(N_t), N_t)$. Thus,

$$\mathbb{E}[\hat{G}'_t(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)] > \mathbb{E}[G'_t(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)],$$

which contradicts inequality (38). Therefore, $\hat{\Delta}_t(N_t) \geq \Delta_t(N_t)$ for any N_t.

Next, we show $\hat{x}_t(N_t) \geq x_t(N_t)$. Note that $\hat{p}_t(N_t) \leq p_t(N_t)$ implies that $\hat{y}_t(N_t) \geq y_t(N_t)$. Thus,

$$\hat{x}_t(N_t) = \hat{y}_t(N_t) + \hat{\Delta}_t(N_t) \geq y_t(N_t) + \Delta_t(N_t) = x_t(N_t).$$

Finally, to complete the induction, we show that $\hat{\pi}'_t(N_t) \geq \pi'_t(N_t)$ for all N_t. By the envelope theorem,

$$\hat{\pi}'_t(N_t) = (\hat{p}_t(N_t) - c)\gamma'(N_t) + (\eta + \theta\gamma'(N_t))\mathbb{E}[\hat{G}'_t(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)],$$

and

$$\pi'_t(N_t) = (p_t(N_t) - c)\gamma'(N_t) + (\eta + \theta\gamma'(N_t))\mathbb{E}[G'_t(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)],$$

If $\hat{p}_t(N_t) = p_t(N_t)$ and $\hat{\Delta}_t(N_t) = \Delta_t(N_t)$, $\hat{\pi}'_t(N_t) \geq \pi'_t(N_t)$ follows immediately from $\gamma'(N_t) \geq 0$ and $\hat{G}'_t(\cdot) \geq G'_t(\cdot)$.

If $\hat{p}_t(N_t) = p_t(N_t)$ and $\hat{\Delta}_t(N_t) > \Delta_t(N_t)$, Lemma 2 yields that $\partial_{\Delta_t}\hat{O}_t(\hat{\Delta}_t(N_t), \hat{p}_t(N_t), N_t) \geq \partial_{\Delta_t}O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}&\beta + \Lambda'(\hat{\Delta}_t(N_t)) + \sigma\mathbb{E}[\hat{G}'_t(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)1_{\{\xi_t \geq \hat{\Delta}_t(N_t)\}}]\\ \geq& \beta + \Lambda'(\Delta_t(N_t)) + \sigma\mathbb{E}[G'_t(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)1_{\{\xi_t \geq \Delta_t(N_t)\}}].\end{aligned}$$

The concavity of $\Lambda(\cdot)$ suggests that $\Lambda'(\hat{\Delta}_t(N_t)) \leq \Lambda'(\Delta_t(N_t))$ and, thus,

$$\begin{aligned}&\mathbb{E}[\hat{G}'_t(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)1_{\{\xi_t \geq \hat{\Delta}_t(N_t)\}}]\\ \geq& \mathbb{E}[G'_t(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)1_{\{\xi_t \geq \Delta_t(N_t)\}}].\end{aligned} \tag{39}$$

Since $\hat{\Delta}_t(N_t) > \Delta_t(N_t)$, it follows immediately that, for any realization of ξ_t,

$$\begin{aligned}&\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)1_{\{\xi_t < \hat{\Delta}_t(N_t)\}}\\ \geq &G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N))^+)1_{\{\xi_t < \Delta_t(N_t)\}}.\end{aligned}$$

Integrate over ξ_t and we have

$$\begin{aligned}&\mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)1_{\{\xi_t < \hat{\Delta}_t(N_t)\}}]\\ \geq &\mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N))^+)1_{\{\xi_t < \Delta_t(N_t)\}}].\end{aligned} \tag{40}$$

Sum up (39) and (40) and we have:

$$\begin{aligned}&\mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)]\\ \geq &\mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N))^+)].\end{aligned} \tag{41}$$

Plugging (41) into the formulas of $\hat{\pi}_t'(\cdot)$ and $\pi_t'(\cdot)$, we have that the inequality $\hat{\pi}_t'(N_t) \geq \pi_t'(N_t)$ follows immediately from $\hat{p}_t(N_t) = p_t(N_t)$.

If $\hat{p}_t(N_t) < p_t(N_t)$, Lemma 2 yields that $\partial_{p_t}\hat{O}_t(\hat{\Delta}_t(N_t), \hat{p}_t(N_t), N_t) \leq \partial_{p_t} O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}&\partial_{p_t} Q_t(\hat{p}_t(N_t), N_t) - \theta\mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)]\\ \leq &\partial_{p_t} Q_t(p_t(N_t), N_t) - \theta\mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)],\end{aligned}$$

i.e.,

$$\begin{aligned}&\bar{V}_t + c - 2\hat{p}_t(N_t) + \gamma(N_t) - \theta\mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)]\\ \leq &\bar{V}_t + c - 2p_t(N_t) + \gamma(N_t) - \theta\mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)].\end{aligned}$$

Thus,

$$\begin{aligned}&(\hat{p}_t(N_t) - p_t(N_t)) + \theta(\mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)]\\ &- \mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)])\\ \geq &\hat{p}_t(N_t) - p_t(N_t)\\ > &0.\end{aligned}$$

Moreover,

$$\begin{aligned}&\mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)]\\ &- \mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)]\\ \geq &\frac{2}{\theta}(p_t(N_t) - \hat{p}_t(N_t))\\ > &0.\end{aligned}$$

Therefore,

$$\begin{aligned}\hat{\pi}_t'(N_t) - \pi_t'(N_t) = &((\hat{p}_t(N_t) - p_t(N_t)) + \theta(\mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)]\\ &- \mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)])\gamma'(N_t)\\ &+ \eta(\mathbb{E}[\hat{G}_t'(\eta N_t + \theta(\hat{y}_t(N_t) + \xi_t) - \sigma(\xi_t - \hat{\Delta}_t(N_t))^+)]\\ &- \mathbb{E}[G_t'(\eta N_t + \theta(y_t(N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)])\\ \geq &0.\end{aligned}$$

Hence, $\hat{\pi}_t'(N_t) \geq \pi_t'(N_t)$ for all N_t. This completes the induction and, thus, the proof of Theorem 6. *Q.E.D.*

Proof of Theorem 7: Part (a). For any $\hat{N}_t > N_t$, since $p_t(\hat{N}_t) \geq p_t(N_t)$ (Theorem 4(a)),

$$y_t(\hat{N}_t) - y_t(N_t) = \gamma(\hat{N}_t) - \gamma(N_t) - (p_t(\hat{N}_t) - p_t(N_t)) \leq \gamma(\hat{N}_t) - \gamma(N_t) \leq (\hat{N}_t - N_t)\gamma'(N_t), \tag{42}$$

where the last inequality follows from the concavity of $\gamma(\cdot)$.

Let $\mathcal{N} := \min\{N_t \geq 0 : \theta\gamma'(N_t) \leq 1-\eta\}$. Since $\lim_{N_t \to +\infty} \gamma'(N_t) = 0$, $\mathcal{N} < +\infty$. Since $I_T \leq x_T(N_T)$, $I_t \leq x_t(N_t)$ for all t with probability 1. Thus,

$$\mathbb{E}[N_{t-1}|N_t] = \eta N_t + \theta y_t(N_t) - \sigma\mathbb{E}(-\Delta_t(N_t) + \xi_t)^+$$

We now show that, if $N_t > \mathcal{N}$, $\mathbb{E}[N_{t-1}|N_t] - N_t$ is decreasing in N_t. For any $\hat{N}_t > N_t > \mathcal{N}$, we have

$$\begin{aligned}
\mathbb{E}[N_{t-1}|\hat{N}_t] - \hat{N}_t - (\mathbb{E}[N_{t-1}|N_t] - N_t) =& \eta\hat{N}_t + \theta y_t(\hat{N}_t) - \sigma\mathbb{E}(\xi_t - \Delta_t(\hat{N}_t))^+ - \hat{N}_t - (\eta N_t + \theta y_t(N_t) - \sigma\mathbb{E}(\xi_t - \Delta_t(N_t))^+ - N_t) \\
=& -(1-\eta)(\hat{N}_t - N_t) + \theta(y_t(\hat{N}_t) - y_t(N_t)) - \sigma(\mathbb{E}(\xi_t - \Delta_t(\hat{N}_t))^+ - \mathbb{E}(\xi_t - \Delta_t(N_t))^+) \\
\leq& -(1-\eta)(\hat{N}_t - N_t) + \theta\gamma'(N_t)((\hat{N}_t - N_t)) \\
<& [(1-\theta) - (1-\theta)]((\hat{N}_t - N_t)) \\
=& 0,
\end{aligned} \tag{43}$$

where the first inequality follows from (42) and $\Delta_t(\hat{N}_t) \leq \Delta_t(N_t)$ (Theorem 4(b)), and the second inequality from the definition of $\mathcal{N}$.

Let $g_t(N_t) := \mathbb{E}[N_{t-1}|N_t] - N_t$. Clearly, $g_t(\cdot)$ is continuous in N_t. On the other hand, $g_t(0) = \mathbb{E}[N_{t-1}|0] = \theta\mathbb{E}(x_t(0) \wedge D_t(p_t(0), 0)) + (\theta - \sigma)\mathbb{E}(D_t(p_t(0), n) - x_t(0))^+$. By Theorem 1(a), $x_t(0) > 0$ and thus $g_t(0) > 0$. Since $\lim_{N_t \to +\infty} \gamma'(N_t) = 0$, when N_t is sufficiently large, $g_t(N_t) \leq -(1-\eta)N_t + C$ for some constant C. Hence, $\lim_{N_t \to +\infty} g_t(N_t) = -\infty$. By the concavity of $\gamma(\cdot)$ and Lemma 7 in Appendix D, it is straightforward to check that, if $g_t(\cdot)$ is decreasing at the point N_t, it is strictly decreasing at any point $\hat{N}_t > N_t$. By (43), there exists a threshold $\bar{\mathcal{N}}_t \leq \mathcal{N}$, such that $g_t(N_t) > 0$ on the region $[0, \bar{\mathcal{N}}_t]$ and it is strictly decreasing in N_t on the region $[\bar{\mathcal{N}}_t, +\infty)$, with $\lim_{N_t \to +\infty} g_t(N_t) = -\infty$. Therefore, there exists a unique $\bar{N}_t > \bar{\mathcal{N}}_t$ such that $g_t(N_t) > 0$ for $N_t < \bar{N}_t$ and $g_t(N_t) < 0$ for $N_t > \bar{N}_t$, i.e., $\mathbb{E}[N_{t-1}|N_t] > N_t$ if $N_t < \bar{N}_t$ and $\mathbb{E}[N_{t-1}|N_t] < N_t$ if $N_t > \bar{N}_t$. Since $g_t(0) > 0$ and $\lim_{N_t \to +\infty} g_t(N_t) = -\infty$, $\bar{N}_t \in (0, +\infty)$.

Part (b). By Theorem 5, $\Delta_t(N)$ is increasing, whereas $p_t(N)$ is decreasing, in t for any $N \geq 0$. Therefore, $\Delta(\cdot) := \lim_{t \to +\infty} \Delta_t(\cdot)$ and $p(\cdot) := \lim_{t \to +\infty} p_t(\cdot)$ is the optimal safety-stock and price in the infinite-horizon discounted reward model ($T = +\infty$). Hence, in this case,

$$N_{t-1} = \eta N_t + \theta(y(N_t) + \xi_t) - \sigma(-\Delta(N_t) + \xi_t)^+,$$

where $y(N_t) = \mathbb{E}[D_t(p(N_t), N_t)] = \bar{V} - p(N_t) + \gamma(N_t)$. By Theorem 11.10 in Stokey et al. (1989), the Markov process $\{N_t : t \in \mathbb{Z}\}$ is ergodic and, thus, has a stationary distribution $\nu(\cdot)$.

Using the same technique as the proof of part (a), we can show that the threshold $\bar{N}$ exists. The convergence result $\bar{N} = \lim_{t \to +\infty} \bar{N}_t$ follows immediately from the monotonicity that $\Delta_t(\cdot)$ is increasing and $p_t(\cdot)$ is decreasing in t.

Part (c). It suffices to show that, as $\eta \uparrow 1$, $\bar{N}_t \uparrow +\infty$ for each t. Since $\mathbb{E}[N_{t-1}|N_t] = \eta N_t + \theta y_t(N_t) - \sigma\mathbb{E}(-\Delta_t(N_t)+\xi_t)^+$ is increasing in η for any N_t. Note that $\bar{N}_t = \inf\{N_t : \mathbb{E}[N_{t-1}|N_t] < N_t\}$, so it follows immediately from the monotonicity of $\mathbb{E}[N_{t-1}|N_t]$ in η that $\bar{N}_t$ is increasing in η. Hence, $\bar{N} = \lim_{t\to+\infty}\bar{N}_t$ is increasing in η as well.

Finally, it remains to show that as $\eta \uparrow 1$, $\bar{N}_t \uparrow +\infty$. Assume, to the contrary, that there exists a uniform upper bound $\Gamma_t \in (0,+\infty)$ such that $\bar{N}_t < \Gamma_t$ for all $\eta < 1$. Hence,

$$g_t(N_t) = -(1-\eta)N_t + y_t(N_t) - \sigma\mathbb{E}(-\Delta_t(N_t)+\xi_t)^+ < 0, \text{ for all } N_t \geq \Gamma_t \text{ and } \eta < 1.$$

Note that $y_t(N_t) - \sigma\mathbb{E}(-\Delta_t(N_t)+\xi_t)^+$ is the average number of customers who join the network in period t. Since $D_t(p_t, N_t) \geq 0$ with probability 1, we have $y_t(N_t) - \sigma\mathbb{E}(-\Delta_t(N_t)+\xi_t)^+ \geq \underline{y}$ for some $\underline{y} > 0$. Therefore, if $\eta > \max\{1-\underline{y}/(2\Gamma_t), 0\}$,

$$g_t(N_t) = -(1-\eta)N_t + y_t(N_t) - \sigma\mathbb{E}(-\Delta_t(N_t)+\xi_t)^+ > -\left(1-\left(1-\frac{\underline{y}}{2\Gamma_t}\right)\right)N_t + \underline{y} = \underline{y}\left(1-\frac{N_t}{2\Gamma_t}\right).$$

Thus, if $N_t = 1.5\Gamma_t > \Gamma_t$, $g_t(N_t) > 0.25\underline{y} > 0$, which contradicts that $g_t(N_t) < 0$ for all $N_t \geq \Gamma_t$. Therefore, as $\eta \uparrow 1$, $\bar{N}_t \uparrow +\infty$ and thus $\bar{N} = \lim_{t\to+\infty}\bar{N}_t$ also increases to infinity as η increases to 1. *Q.E.D.*

Proof of Theorem 8: Part (a). First $v_t^w(\cdot,\cdot) \leq v_t(\cdot,\cdot)$ follows immediately from that $v_t(\cdot,\cdot)$ is the optimal expected profit in periods $t, t-1, \cdots, 1$.

If $w \geq t$, the length of the moving time-window exceeds the total planning horizon length. Therefore, if $w \geq t$, the w-heuristic is the optimal policy and, hence, $v_t^w(\cdot,\cdot) = v_t(\cdot,\cdot)$.

It remains to show that if $w \leq t-2$, $v_t^w(\cdot,\cdot) \leq v_t^{w+1}(\cdot,\cdot)$. Note that

$$\begin{aligned}
v_t(I_t, N_t) =& \mathbb{E}[\sum_{1\leq\tau\leq t}\alpha^{t-\tau}\{Q_\tau(p_\tau, N_\tau) + \beta\Delta_\tau + \Lambda(\Delta_\tau) + cI_\tau\} + \alpha^t I_0 | I_t, N_t], \\
&s.t. \text{ for each } 1 \leq \tau \leq t: \\
p_\tau =& p_\tau(N_\tau); \\
\Delta_\tau =& \Delta_\tau(N_\tau); \\
N_{\tau-1} =& \eta N_\tau + \theta D_\tau(p_\tau, N_\tau) - \sigma(\xi_\tau - \Delta_\tau(p_\tau, N_\tau))^+ + \epsilon_\tau; \\
I_{\tau-1} =& \bar{V}_\tau + \gamma(N_\tau) - p_\tau + \Delta_\tau(N_\tau) - D_\tau(p_\tau, N_\tau).
\end{aligned}$$

Analogously, we have

$$
\begin{aligned}
v_t^w(I_t,N_t) =& \mathbb{E}[\sum_{1\le\tau\le t}\alpha^{t-\tau}\{Q_\tau(p_\tau,N_\tau)+\beta\Delta_\tau+\Lambda(\Delta_\tau)+cI_\tau\}+\alpha^t I_0|I_t,N_t],\\
s.t.\ & \text{for each } 1\le\tau\le w:\\
& p_\tau = p_\tau(N_\tau);\\
& \Delta_\tau = \Delta_\tau(N_\tau);\\
& \text{for each } w+1\le\tau\le t:\\
& p_\tau = p_w(N_\tau);\\
& \Delta_\tau = \Delta_w(N_\tau);\\
& \text{for each } 1\le\tau\le t:\\
& N_{\tau-1} = \eta N_\tau+\theta D_\tau(p_\tau,N_\tau)-\sigma(\xi_\tau-\Delta_\tau(p_\tau,N_\tau))^+ +\epsilon_\tau;\\
& I_{\tau-1} = \bar{V}_\tau+\gamma(N_\tau)-p_\tau+\Delta_\tau(N_\tau)-D_\tau(p_\tau,N_\tau);
\end{aligned}
$$

and

$$
\begin{aligned}
v_t^{w+1}(I_t,N_t) =& \mathbb{E}[\sum_{1\le\tau\le t}\alpha^{t-\tau}\{Q_\tau(p_\tau,N_\tau)+\beta\Delta_\tau+\Lambda(\Delta_\tau)+cI_\tau\}+\alpha^t I_0|I_t,N_t],\\
s.t.\ & \text{for each } 1\le\tau\le w+1:\\
& p_\tau = p_\tau(N_\tau);\\
& \Delta_\tau = \Delta_\tau(N_\tau);\\
& \text{for each } w+2\le\tau\le t:\\
& p_\tau = p_w(N_\tau);\\
& \Delta_\tau = \Delta_w(N_\tau);\\
& \text{for each } 1\le\tau\le t:\\
& N_{\tau-1} = \eta N_\tau+\theta D_\tau(p_\tau,N_\tau)-\sigma(\xi_\tau-\Delta_\tau(p_\tau,N_\tau))^+ +\epsilon_\tau;\\
& I_{\tau-1} = \bar{V}_\tau+\gamma(N_\tau)-p_\tau+\Delta_\tau(N_\tau)-D_\tau(p_\tau,N_\tau);
\end{aligned}
$$

Since $I_T\le x_T(N_t)$, Theorem 5 implies that $\Delta_\tau(\cdot)\ge\Delta_{w+1}(\cdot)\ge\Delta_w(\cdot)$ and $p_\tau(\cdot)\le p_{w+1}(\cdot)\le p_w(\cdot)$ for all $\tau\ge w+1$. Moreover, $(\Delta_\tau(\cdot),p_\tau(\cdot))$ is the optimal policy for the dynamic program. Putting everything together, it follows immediately that $v_t^w(\cdot,\cdot)\le v_t^{w+1}(\cdot,\cdot)\le v_t(\cdot,\cdot)$ for all $t\ge w+2$. This finishes the proof of part (a).

Part (b). The inequality $v^w(\cdot,\cdot)\le v^{w+1}(\cdot,\cdot)\le v(\cdot,\cdot)$ follows from $v_t^w(\cdot,\cdot)\le v_t^{w+1}(\cdot,\cdot)\le v_t(\cdot,\cdot)$ by letting T approach infinity. Note that $v_w(\cdot,\cdot)\le v^w(\cdot,\cdot)\le v(\cdot,\cdot)$. Hence, $\sup|v^w(\cdot,\cdot)-v(\cdot,\cdot)|\le\sup|v_w(\cdot,\cdot)-v(\cdot,\cdot)|$. Let $\mathcal{T}$ be the operator acted on a concave and continuously differentiable function $f(\cdot,\cdot)$ that satisfies

$$
\begin{aligned}
\mathcal{T}[f(I,N)] &= \max_{(x,p)\in\hat{\mathcal{F}}(I)}\{R(p,N)+\beta x+\Lambda(x-y(p,N))\\
&\quad +\mathbb{E}[\Psi(x_t-y(p,N)-\xi,\eta N+\theta(y(p,N)+\xi)-\sigma(y(p,N)+\xi-x)^+)]\},\\
\text{with } \Psi(x,y) &:= \alpha\mathbb{E}\{[f(x,y+\epsilon_t)-cx]\},\\
R(p,N) &:= (p-\alpha c-b)(\bar{V}+\gamma(N)-p),\\
y(p,N) &:= \bar{V}+\gamma(N)-p.
\end{aligned}
$$

By Theorem 9.6 in Stokey et al. (1989), $\mathcal{T}$ is a contraction mapping with contraction factor α under the sup norm. Since $v(\cdot,\cdot)$ is the fixed point of $\mathcal{T}$ and $v_w = \mathcal{T}^w[v_0(\cdot,\cdot)]$, we have

$$\sup|v_w(\cdot,\cdot) - v(\cdot,\cdot)| \leq \alpha^w \sup|v_0(\cdot,\cdot) - v(\cdot,\cdot)|.$$

Let $C := \sup|v_0(\cdot,\cdot) - v(\cdot,\cdot)| > 0$ for a given initial state (I,N), and $\delta := -\log(\alpha) > 0$. Thus, we have

$$\sup|v^w(\cdot,\cdot) - v(\cdot,\cdot)| \leq \sup|v_w(\cdot,\cdot) - v(\cdot,\cdot)| \leq C\mathrm{e}^{-\delta w}. \tag{44}$$

By (44), $\lim_{w\to+\infty} v^w(\cdot,\cdot) = v(\cdot,\cdot)$ follows immediately for any initial state (I,N). This proves part (b). *Q.E.D.*

Proof of Lemma 5: Part (a). Part (a) follows from the same argument as the proof of Lemma 3, so we omit its proof for brevity.

Part (b). The optimal value function $v_t^e(I_t,N_t)$ satisfies the following recursive scheme:

$$v_t^e(I_t,N_t) = cI_t + \max_{(x_t,p_t,n_t)\in\hat{\mathcal{F}}_e(I_t)} J_t^e(x_t,p_t,n_t,N_t), \tag{45}$$

where $\hat{\mathcal{F}}_e(I_t) := [I_t,+\infty)\times[\underline{p},\bar{p}]\times[0,+\infty)$ denotes the set of feasible decisions and

$$\begin{aligned} J_t^e(x_t,p_t,n_t,N_t) =& R_t(p_t,N_t) + \beta x_t + \Lambda(x_t - y_t(p_t,N_t)) - c_n(n_t) \\ &+ \mathbb{E}[\Psi_t^e(x_t - y_t(p_t,N_t) - \xi_t, \eta N_t + \theta(y_t(p_t,N_t)+\xi_t) - \sigma(y_t(p_t,N_t)+\xi_t - x_t)^+ + n_t)], \end{aligned} \tag{46}$$

with $\Psi_t^e(x,y) := \alpha\mathbb{E}\{v_{t-1}^e(x, y+\epsilon_t) - cx\}$.

The derivation of (46) is given as follows:

$$\begin{aligned} J_t^e(x_t,p_t,n_t,N_t) :=& -cI_t + \mathbb{E}\{p_tD_t(p_t,N_t) - c(x_t - I_t) - h(x_t - D_t(p_t,N_t))^+ - b(x_t - D_t(p_t,N_t))^- - c_n(n_t) \\ & +\alpha v_{t-1}^e(x_t - D_t(p_t,N_t), \theta D_t(p_t,N_t) + \eta N_t - \sigma(D_t(p_t,N_t) - x_t)^+ + n_t + \epsilon_t)|N_t\}, \\ =& (p_t - \alpha c - b)y_t(p_t,N_t) + (b-(1-\alpha)c)x_t - c_n(n_t) + \mathbb{E}\{-(h+b)(x_t - y_t(p_t,N_t) - \xi_t)^+ \\ & +\alpha[v_{t-1}^e(x_t - y_t(p_t,N_t) - \xi_t, \eta N_t + \theta(y_t(p_t,N_t)+\xi_t) - \sigma(y_t(p_t,N_t)+\xi_t - x_t)^+ + n_t + \epsilon_t) \\ & -c(x_t - y_t(p_t,N_t) - \xi_t)]|N_t\} \\ =& R_t(p_t,N_t) + \beta x_t + \Lambda(x_t - y_t(p_t,N_t)) - c_n(n_t) \\ & +\mathbb{E}[\Psi_t^e(x_t - y_t(p_t,N_t) - \xi_t, \eta N_t + \theta(y_t(p_t,N_t)+\xi_t) - \sigma(y_t(p_t,N_t)+\xi_t - x_t)^+ + n_t)]. \end{aligned}$$

We use $(\hat{x}_t^e(N_t), \hat{p}_t^e(N_t), \hat{n}_t(N_t))$ as the unconstrained optimizer of (46). The same argument as the proof of Lemma 3 yields that $J_t^e(\cdot,\cdot,\cdot,\cdot)$ is jointly concave in (x_t,p_t,n_t,N_t). Hence, if $I_t \leq \hat{x}_t^e(N_t)$, $(x_t^{e*}(I_t,N_t), p_t^{e*}(I_t,N_t), n_t^*(I_t,N_t)) = (\hat{x}_t^e(N_t), \hat{p}_t^e(N_t), \hat{n}_t(N_t))$; otherwise $(I_t > \hat{x}_t^e(N_t))$ $x_t^{e*}(I_t,N_t) = I_t$.

The same argument as the proof of Lemma 1 implies that $\mathbb{P}[\hat{x}_t^e(N_t) - D_t(\hat{p}_t^e(N_t),N_t) \leq \hat{x}_{t-1}^e(N_{t-1})|N_t] = 1$. Hence, the same argument as the proof of Lemma 4 and its discussions enables us to transform the objective function from $J_t^e(x_t,p_t,n_t,N_t)$ to $O_t^e(\Delta_t,p_t,n_t,N_t)$ by letting $\Delta_t = x_t - \mathbb{E}[D_t(p_t,N_t)] = x_t - y_t(p_t,N_t)$, and we have that $(\hat{x}_t^e(N_t), \hat{p}_t^e(N_t), \hat{n}_t(N_t)) = (x_t^e(N_t), p_t^e(N_t), n_t(N_t))$ for all $N_t \geq 0$, where $x_t^e(N_t) = \Delta_t^e(N_t) + y_t(p_t^e(N_t),N_t)$. Hence, as in Theorem 2, if $I_T \leq x_T^e(N_T)$, $I_t \leq x_t^e(N_t)$ for all t with probability 1. Hence, part (b) follows. *Q.E.D.*

Proof of Theorem 9: Part (a). We first show that if (10) holds, $n_t^*(I_t, N) > 0$ for all I_t. Observe that, since $\partial_y \Psi_{t-1}^e(x, y) \geq 0$,

$$\partial_{N_{t-1}} v_{t-1}^e(I_{t-1}, N_{t-1}) \geq (\underline{p} - b - \alpha c)\gamma'(N_{t-1}) - \gamma'(N_{t-1})\Lambda'(\Delta_{t-1}^*),$$

where $\Delta_{t-1}^* := x_{t-1}^{e*}(I_{t-1}, N_{t-1}) - y_{t-1}(I_{t-1}, N_{t-1})$. The first-order condition with respect to x_{t-1} yields that $\Lambda'(\Delta_{t-1}^*) \leq -\beta$ for any realization of ξ_t and ϵ_t. Thus, for any realization of ξ_t and ϵ_t,

$$\partial_{N_{t-1}} v_{t-1}^e(I_{t-1}, N_{t-1}) \geq (\underline{p} - c)\gamma'(N_{t-1}). \tag{47}$$

Therefore, for any Δ_t and $p_t \in [\underline{p}, \bar{p}]$,

$$\begin{aligned}
\partial_{n_t} O_t^e(\Delta_t, p_t, 0, N) \geq & \alpha \mathbb{E}\{\partial_{N_{t-1}} v_{t-1}^e(x_t - D_t(p_t, N), N_{t-1}) | N_t = N\} - c_n'(0) \\
\geq & \alpha \mathbb{E}\{(\underline{p} - c)\gamma'(N_{t-1}) | N_t = N\} - c_n'(0) \\
\geq & \alpha(1 - \iota)(\underline{p} - c)\gamma'(\bar{S}(N)) - c_n'(0) \\
> & 0,
\end{aligned} \tag{48}$$

where the second inequality follows from (47), and the fourth from the assumption (10). The third inequality of (48) follows from the following inequality:

$$\begin{aligned}
\alpha \mathbb{E}[(\underline{p} - c)\gamma'(N_{t-1}) | N_t = N] = & \ \alpha \mathbb{E}_{N_{t-1} \geq \bar{S}(N)}[(\underline{p} - c)\gamma'(N_{t-1}) | N_t = N] \\
& + \alpha \mathbb{E}_{N_{t-1} < \bar{S}(N)}[(\underline{p} - c)\gamma'(N_{t-1}) | N_t = N] \\
\geq & \ 0 + \alpha \mathbb{E}_{N_{t-1} < \bar{S}(N)}[(\underline{p} - c)\gamma'(\bar{S}(N))] \\
\geq & \ \alpha(1 - \iota)(\underline{p} - c)\gamma'(\bar{S}(N)),
\end{aligned}$$

where the first inequality follows from the concavity of $\gamma(\cdot)$, and the second from the definition of $\bar{S}(N)$. The inequality (48) yields that $n_t^*(I_t, N) > 0$ for all I_t.

Since $\gamma(\cdot)$ is continuously increasing in N_t, $\bar{S}(N)$ is continuously increasing in N. The concavity of $\gamma(\cdot)$ implies that $\gamma'(\bar{S}(N))$ is continuously decreasing in N. Therefore, let

$$N^*(\iota) := \max\{N \geq 0 : \alpha(1 - \iota)[(\underline{p} - c)\gamma'(\bar{S}(N))] > c_n'(0)\}.$$

We have (10) holds for all $N < N^*(\iota)$. This completes the proof of part (a).

Part (b). Since $\gamma(\cdot)$ is concavely increasing in N_t,

$$\partial_{N_{t-1}} v_{t-1}^e(I_{t-1}, N_{t-1}) \leq \partial_{N_{t-1}} v_t^e(I_{t-1}, 0) \leq (\sum_{\tau=1}^{t-1} (\alpha\eta)^{\tau-1})(\bar{p} - c)\gamma'(0).$$

Thus, if $\alpha(\sum_{\tau=1}^{t-1} (\alpha\eta)^{\tau-1})(\bar{p} - c)\gamma'(0) \leq c_n'(0)$, then for any (x_t, p_t, n_t, N_t),

$$\begin{aligned}
\partial_{n_t} J_t^e(x_t, p_t, n_t, N_t) \leq & \alpha \mathbb{E}\{\partial_{N_{t-1}} v_{t-1}^e(x_t - D_t(p_t, N_t), N_{t-1} + n_t) | N_t = N\} - c_n'(0) \\
\leq & \alpha(\sum_{\tau=1}^{t-1} (\alpha\eta)^{\tau-1})(\bar{p} - c)\gamma'(0) - c_n'(0) \\
\leq & 0.
\end{aligned}$$

Hence, $n_t^*(I_t, N_t) = 0$ for all (I_t, N_t). This completes the proof of part (b). *Q.E.D.*

Proof of Theorem 10: Parts (a)-(c). We prove parts (a)-(c) together by backward induction. More specifically, we show that if $\partial_{N_{t-1}}\pi_{t-1}^e(\cdot) \le \partial_{N_{t-1}}\pi_{t-1}(\cdot)$ for all $N_{t-1} \ge 0$, (i) $p_t^e(N_t) \ge p_t(N_t)$, (ii) $\Delta_t^e(N_t) \le \Delta_t(N_t)$, (iii) $x_t^e(N_t) \le x_t(N_t)$, and (iv) $\partial_{N_t}\pi_t^e(\cdot) \le \partial_{N_t}\pi_t(\cdot)$ for all $N_t \ge 0$. Since $\partial_{N_0}\pi_0^e(\cdot) = \partial_{N_0}\pi_0(\cdot) \equiv 0$, the initial condition is satisfied. Note that $\partial_{N_{t-1}}\pi_{t-1}^e(N_{t-1}) \le \partial_{N_{t-1}}\pi_{t-1}(N_{t-1})$ for all $N_{t-1} \ge 0$ implies that

$$\partial_y G_t^e(y) = \alpha\mathbb{E}\{\partial_{N_{t-1}}\pi_{t-1}^e(y+\epsilon_t)\} \le \alpha\mathbb{E}\{\partial_{N_{t-1}}\pi_{t-1}(y+\epsilon_t)\} = \partial_y G_t(y),$$

for all y.

We first show that $p_t^e(N_t) \ge p_t(N_t)$ for all N_t. Assume, to the contrary, that $p_t^e(N_t) < p_t(N_t)$ for some N_t. Lemma 2 implies that $\partial_{p_t}O_t^e(\Delta_t^e(N_t), p_t^e(N_t), n_t(N_t), N_t) \le \partial_{p_t}O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}&\partial_{p_t}Q(p_t^e(N_t), N_t) - \theta\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)]\\ \le&\partial_{p_t}Q_t(p_t(N_t), N_t) - \theta\mathbb{E}[\partial y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)].\end{aligned}$$

Since $Q_t(\cdot, N_t)$ is strictly concave in p_t and $p_t^e(N_t) < p_t(N_t)$, $\partial_{p_t}Q(p_t^e(N_t), N_t) > \partial_{p_t}Q_t(p_t(N_t), N_t)$. Thus,

$$\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)] > \mathbb{E}[\partial y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)]. \tag{49}$$

Note that $\partial_y G_t^e(\cdot) \le \partial_y G_t(\cdot)$ for all y, $p_t^e(N_t) < p_t(N_t)$, and $n_t(N_t) \ge 0$. Thus, the concavity of $G_t^e(\cdot)$ and $G_t(\cdot)$, together with the inequality (49), implies that $\sigma > 0$ and $\Delta_t^e(N_t) < \Delta_t(N_t)$. Thus, Lemma 2 implies that $\partial_{\Delta_t}O_t^e(\Delta_t^e(N_t), p_t^e(N_t), n_t(N_t), N_t) \le \partial_{\Delta_t}O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}&\beta + \Lambda'(\Delta_t^e(N_t)) + \sigma\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)1_{\{\xi_t \ge \Delta_t^e(N_t)\}}]\\ \le&\beta + \Lambda'(\Delta_t(N_t)) + \sigma\mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)1_{\{\xi_t \ge \Delta_t(N_t)\}}].\end{aligned}$$

The concavity of $\Lambda(\cdot)$ suggests that $\Lambda'(\Delta_t^e(N_t)) \ge \Lambda'(\Delta_t(N_t))$ and, thus,

$$\begin{aligned}&\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)1_{\{\xi_t \ge \Delta_t^e(N_t)\}}]\\ \le&\mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)1_{\{\xi_t \ge \Delta_t(N_t)\}}].\end{aligned} \tag{50}$$

Since $\Delta_t^e(N_t) < \Delta_t(N_t)$ and $0 \le \partial y G_t^e(\cdot) \le \partial_y G_t(\cdot)$, it follows immediately that, for any realization of ξ_t,

$$\begin{aligned}&\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)1_{\{\xi_t < \Delta_t^e(N_t)\}}\\ \le&\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)1_{\{\xi_t < \Delta_t(N_t)\}}.\end{aligned}$$

Integrate over ξ_t and we have

$$\begin{aligned}&\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)1_{\{\xi_t < \Delta_t^e(N_t)\}}]\\ \le&\mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)1_{\{\xi_t < \Delta_t(N_t)\}}].\end{aligned} \tag{51}$$

Sum up (50) and (51) and we have:

$$\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)] \le \mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)],$$

which contradicts (49). Therefore, $p_t^e(N_t) \ge p_t(N_t)$ for all N_t.

Next, we show that $\Delta_t^e(N_t) \ge \Delta_t(N_t)$. If $\sigma = 0$, it is straightforward to show that $\Delta_t^e(N_t) = \Delta_t(N_t) = \Delta_*$. Hence, we restrict ourselves to the interesting case of $\sigma > 0$.

Assume, to the contrary, that $\Delta_t^e(N_t) > \Delta_t(N_t)$. Lemma 2 implies that $\partial_{\Delta_t} O_t^e(\Delta_t^e(N_t), p_t^e(N_t), n_t(N_t), N_t) \geq \partial_{\Delta_t} O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}&\beta + \Lambda'(\Delta_t^e(N_t)) + \sigma\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+) 1_{\{\xi_t \geq \Delta_t^e(N_t)\}}]\\ \geq &\beta + \Lambda'(\Delta_t(N_t)) + \sigma\mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+) 1_{\{\xi_t \geq \Delta_t(N_t)\}}].\end{aligned}$$

The concavity of $\Lambda(\cdot)$ suggests that $\Lambda'(\Delta_t^e(N_t)) \leq \Lambda'(\Delta_t(N_t))$ and, thus,

$$\begin{aligned}&\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+) 1_{\{\xi_t \geq \Delta_t^e(N_t)\}}]\\ \geq &\mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+) 1_{\{\xi_t \geq \Delta_t(N_t)\}}].\end{aligned} \tag{52}$$

The concavity of $G_t^e(\cdot)$ and $G_t(\cdot)$ and that $n_t(N_t) \geq 0$ and $\partial_y G_t^e(\cdot) \leq \partial_y G_t(\cdot)$ imply that $y_t(p_t^e(N_t), N_t) < y_t(p_t(N_t), N_t)$ and, thus, $p_t^e(N_t) > p_t(N_t)$. Since $\Delta_t^e(N_t) < \Delta_t(N_t)$, it follows immediately that, for any realization of ξ_t,

$$\begin{aligned}&\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+) 1_{\{\xi_t < \Delta_t^e(N_t)\}}\\ \geq &\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+) 1_{\{\xi_t < \Delta_t(N_t)\}}.\end{aligned}$$

Integrate over ξ_t and we have

$$\begin{aligned}&\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+) 1_{\{\xi_t < \Delta_t^e(N_t)\}}]\\ \geq &\mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+) 1_{\{\xi_t < \Delta_t(N_t)\}}].\end{aligned} \tag{53}$$

Sum up (52) and (53) and we have:

$$\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)] \geq \mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)]. \tag{54}$$

By Lemma 2, $p_t^e(N_t) > p_t(N_t)$ yields that $\partial_{p_t} O_t^e(\Delta_t^e(N_t), p_t^e(N_t), n_t(N_t), N_t) \geq \partial_{p_t} O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}&\partial_{p_t} Q(p_t^e(N_t), N_t) - \theta\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)]\\ \geq &\partial_{p_t} Q_t(p_t(N_t), N_t) - \theta\mathbb{E}[\partial y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)].\end{aligned}$$

Since $Q_t(\cdot, N_t)$ is strictly concave in p_t, $\partial_{p_t} Q(p_t^e(N_t), N_t) < \partial_{p_t} Q_t(p_t(N_t), N_t)$. Thus,

$$\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)] < \mathbb{E}[\partial y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)],$$

which contradicts inequality (54). Therefore, $\Delta_t^e(N_t) \leq \Delta_t(N_t)$ for any N_t.

Next, we show $x_t^e(N_t) \leq x_t(N_t)$. Note that $p_t^e(N_t) \geq p_t(N_t)$ implies that $y_t(p_t^e(N_t), N_t) \leq y_t(p_t(N_t), N_t)$. Thus,

$$x_t^e(N_t) = y_t(p_t^e(N_t), N_t) + \Delta_t^e(N_t) \leq y_t(p_t(N_t), N_t) + \Delta_t(N_t) = x_t(N_t).$$

Finally, to complete the induction, we show that $\partial_{N_t} \pi_t^e(N_t) \geq \partial_{N_t} \pi_t(N_t)$ for all $N_t \geq 0$. By the envelope theorem,

$$\partial_{N_t} \pi_t^e(N_t) = (p_t^e(N_t) - c)\gamma'(N_t) + (\eta + \theta\gamma'(N_t))\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)],$$

and

$$\partial_{N_t} \pi_t(N_t) = (p_t(N_t) - c)\gamma'(N_t) + (\eta + \theta\gamma'(N_t))\mathbb{E}[\partial y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)].$$

If $p_t^e(N_t) = p_t(N_t)$ and $\Delta_t^e(N_t) = \Delta_t(N_t)$, $\partial_{N_t} \pi_t^e(N_t) \leq \partial_{N_t} \pi_t(N_t)$ follows immediately from $\gamma'(N) \geq 0$ and $\partial_y G_t^e(\cdot) \leq \partial_y G_t(\cdot)$ for all y.

If $p_t^e(N_t) = p_t(N_t)$ and $\Delta_t^e(N_t) < \Delta_t(N_t)$, Lemma 2 yields that $\partial_{\Delta_t} O_t^e(\Delta_t^e(N_t), p_t^e(N_t), n_t(N_t), N_t) \leq \partial_{\Delta_t} O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}&\beta + \Lambda'(\Delta_t^e(N_t)) + \sigma\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+) 1_{\{\xi_t \geq \Delta_t^e(N_t)\}}]\\ \leq&\beta + \Lambda'(\Delta_t(N_t)) + \sigma\mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+) 1_{\{\xi_t \geq \Delta_t(N_t)\}}].\end{aligned}$$

The concavity of $\Lambda(\cdot)$ suggests that $\Lambda'(\Delta_t^e(N_t)) \geq \Lambda'(\Delta_t(N_t))$ and, thus,

$$\begin{aligned}&\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+) 1_{\{\xi_t \geq \Delta_t^e(N_t)\}}]\\ \leq&\mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+) 1_{\{\xi_t \geq \Delta_t(N_t)\}}].\end{aligned} \tag{55}$$

Since $\Delta_t^e(N_t) < \Delta_t(N_t)$, it follows immediately that, for any realization of ξ_t,

$$\begin{aligned}&\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+) 1_{\{\xi_t < \Delta_t^e(N_t)\}}\\ \leq&\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+) 1_{\{\xi_t < \Delta_t(N_t)\}}.\end{aligned}$$

Integrate over ξ_t and we have

$$\begin{aligned}&\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+) 1_{\{\xi_t < \Delta_t^e(N_t)\}}]\\ \leq&\mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+) 1_{\{\xi_t < \Delta_t(N_t)\}}].\end{aligned} \tag{56}$$

Sum up (55) and (56) and we have:

$$\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)] \leq \mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)]. \tag{57}$$

Plugging (57) into the formulas of $\partial_{N_t} \pi_t^e(\cdot)$ and $\partial_{N_t} \pi_t(\cdot)$, we have the inequality $\partial_{N_t} \pi_t^e(N_t) \leq \partial_{N_t} \pi_t(N_t)$ follows immediately from $p_t^e(N) = p_t(N_t)$.

If $p_t^e(N_t) > p_t(N_t)$, Lemma 2 yields that $\partial_{p_t} O_t^e(\Delta_t^e(N_t), p_t^e(N_t), n_t(N_t), N_t) \geq \partial_{p_t} O_t(\Delta_t(N_t), p_t(N_t), N_t)$, i.e.,

$$\begin{aligned}&\partial_{p_t} Q_t(p_t^e(N_t), N_t) - \theta\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)]\\ \geq&\partial_{p_t} Q_t(p_t(N_t), N_t) - \theta\mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)],\end{aligned}$$

i.e.,

$$\begin{aligned}&\bar{V}_t + c - 2p_t^e(N_t) + \gamma(N_t) - \theta\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)]\\ \leq&\bar{V}_t + c - 2p_t(N_t) + \gamma(N_t) - \theta\mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)].\end{aligned}$$

Thus,

$$\begin{aligned}&(p_t^e(N_t) - p_t(N_t)) + \theta(\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)]\\ &- \mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)])\\ \geq&p_t^e(N_t) - p_t(N_t)\\ >&0.\end{aligned}$$

Moreover,

$$\begin{aligned}&\mathbb{E}[\partial_y G_t^e(\eta N_t + \theta(y_t(p_t^e(N_t), N_t) + \xi_t) + n_t(N_t) - \sigma(\xi_t - \Delta_t^e(N_t))^+)]\\ &- \mathbb{E}[\partial_y G_t(\eta N_t + \theta(y_t(p_t(N_t), N_t) + \xi_t) - \sigma(\xi_t - \Delta_t(N_t))^+)]\\ \geq&\frac{2}{\theta}(p_t^e(N_t) - p_t(N_t))\\ >&0.\end{aligned}$$

Therefore,

$$\begin{aligned}\partial_{N_t}\pi_t^e(N_t)-\partial_{N_t}\pi_t(N_t)=&((p_t^e(N_t)-p_t(N_t))+\theta(\mathbb{E}[\partial_y G_t^e(\eta N_t+\theta(y_t(p_t^e(N_t),N_t)+\xi_t)+n_t(N_t)-\sigma(\xi_t-\Delta_t^e(N_t))^+)]\\&-\mathbb{E}[\partial_y G_t(\eta N_t+\theta(y_t(p_t(N_t),N_t)+\xi_t)-\sigma(\xi_t-\Delta_t(N_t))^+)])\gamma'(N)\\&+\eta(\mathbb{E}[\partial_y G_t^e(\eta N_t+\theta(y_t(p_t^e(N_t),N_t)+\xi_t)+n_t(N_t)-\sigma(\xi_t-\Delta_t^e(N_t))^+)]\\&-\mathbb{E}[\partial_y G_t(\eta N_t+\theta(y_t(p_t(N_t),N_t)+\xi_t)-\sigma(\xi_t-\Delta_t(N_t))^+)])\\\ge&0.\end{aligned}$$

Hence, $\partial_{N_t}\pi_t^e(N_t)\ge\partial_{N_t}\pi_t(N_t)$ for all N_t. This completes the induction and, thus, the proof of parts (a)-(c).

Part (d). Note that $\pi_t(\cdot)$ is the normalized optimal profit with the Bellman equation (9) and feasible decision set $\{(x_t,p_t,n_t):\Delta_t\in\mathbb{R},p_t\in[\underline{p},\bar{p}],n_t=0\}\subset\mathcal{F}_e$, which is the feasible decision set associated with the profit $\pi_t^e(\cdot)$. Thus, $\pi_t^e(N_t)\ge\pi_t(N_t)$ for all t and any $N_t\ge0$. If $n_t(N_t)>0$, we must have $\pi_t^e(N_t)>\pi_t(N_t)$. Otherwise there are two lexicographically different policies (one with $n_t(N_t)=0$ and the other with $n_t(N_t)>0$) that generate the same optimal normalized profit $\pi_t(N_t)$. This contradicts the assumption that the lexicographically smallest policy is selected. Thus, $\pi_t^e(N_t)>\pi_t(N_t)$, which establishes part (d). *Q.E.D.*

Proof of Theorem 11: Part (a). We show part (a) by backward induction. More specifically, we show that if $\sigma=\eta=0$ and $v_{t-1}(\cdot,\cdot)$ is supermodular in (I_{t-1},N_{t-1}), $v_t(\cdot,\cdot)$ is supermodular in (I_t,N_t). Since $v_0(I_0,N_0)=cI_0$, the initial condition is satisfied.

Since supermodularity is preserved under expectation, $\bar{\Psi}_t(x,y):=\alpha\mathbb{E}\{[v_{t-1}(x-\xi_t,y+\theta\xi_t+\epsilon_t)-cx]\}$ is supermodular in (x,y). Let $y_t=\bar{V}_t-p_t+\gamma(N_t)$. Observe that

$$\begin{aligned}J_t(x_t,p_t,N_t)=&R_t(p_t,N_t)+\beta x_t+\Lambda(x_t-\bar{V}_t+p_t-\gamma(N_t))+\bar{\Psi}_t(x_t-\bar{V}_t+p_t-\gamma(N_t),\theta(\bar{V}_t-p_t+\gamma(N_t)))\\=&(\bar{V}_t+\gamma(N_t)-y_t-\alpha c-b)y_t+\beta x_t+\Lambda(x_t-y_t)+\bar{\Psi}_t(x_t-y_t,\theta y_t).\end{aligned}$$

Hence,

$$v_t(I_t,N_t)=cI_t+\max_{(x_t,y_t)\in\mathcal{F}_t'(I_t,N_t)}\{(\bar{V}_t+\gamma(N_t)-y_t-\alpha c-b)y_t+\beta x_t+\Lambda(x_t-y_t)+\bar{\Psi}_t(x_t-y_t,\theta y_t)\},$$

where $\mathcal{F}_t'(I_t,N_t):=\{(x_t,y_t):x_t\ge I_t,y_t\in[\bar{V}_t+\gamma(N_t)-\bar{p},\bar{V}_t+\gamma(N_t)-\underline{p}]\}$. Because $\gamma(\cdot)$ is increasing in N_t, $\Lambda(\cdot)$ is concave, and $\bar{\Psi}_t(\cdot,\cdot)$ is concave and supermodular, $(\bar{V}_t+\gamma(N_t)-y_t-\alpha c-b)y_t+\beta x_t+\Lambda(x_t-y_t)+\bar{\Psi}_t(x_t-y_t,\theta y_t)$ is supermodular in (x_t,y_t,N_t). Moreover, it's straightforward to verify that the feasible set $\{(x_t,y_t,I_t,N_t):N_t\ge0,(x_t,y_t)\in\mathcal{F}_t'(I_t)\}$ is a lattice in $\mathbb{R}^4$. Therefore, $v_t(I_t,N_t)$ is supermodular in (I_t,N_t). This completes the induction and, thus, the proof of part (a).

Part (b). The continuity results in parts (b)-(e) all follow from the joint concavity and continuous differentiability of $J_t(\cdot,\cdot,\cdot)$ in (x_t,p_t,N_t). Since $x_t^*(I_t,N_t)=\max\{I_t,x_t(N_t)\}$, $x_t^*(I_t,N_t)$ is increasing in I_t. Moreover, because the objective function $(\bar{V}_t+\gamma(N_t)-y_t-\alpha c-b)y_t+\beta x_t+\Lambda(x_t-y_t)+\bar{\Psi}_t(x_t-y_t,\theta y_t)$ is supermodular in (x_t,y_t,N_t), $x_t^*(I_t,N_t)$ is increasing in N_t as well. This proves part (b).

Part (c). If $I_t\le x_t(N_t)$, $p_t^*(I_t,N_t)=p_t(N_t)$, which is independent of I_t. If $I_t>x_t(N_t)$, $x_t^*(I_t,N_t)=I_t$ and, thus,

$$J_t(x_t^*(I_t,N_t),p_t,N_t)=R_t(p_t,N_t)+\beta I_t+\Lambda(I_t-\bar{V}_t+p_t-\gamma(N_t))+\bar{\Psi}_t(I_t-\bar{V}_t+p_t-\gamma(N_t),\theta(\bar{V}_t-p_t+\gamma(N_t))).\tag{58}$$

Since $\Lambda(\cdot)$ is concave and $\bar{\Psi}_t(\cdot,\cdot)$ is concave and supermodular, $J_t(x_t^*(I_t,N_t),p_t,N_t)$ is submodular in (I_t,p_t). Hence, $p_t^*(I_t,N_t)$ is decreasing in I_t for all (I_t,N_t). By Theorem 4(d), if $I_t \le x_t(N_t)$, $p_t^*(I_t,N_t)=p_t(N_t)$ is increasing in N_t. If $I_t > x_t(N_t)$, we observe from (58) that $J_t(x_t^*(I_t,N_t),p_t,N_t)$ is supermodular in (p_t,N_t). Hence, $p_t^*(I_t,N_t)$ is increasing in N_t for all (I_t,N_t). This proves part (c).

Part (d). If $I_t \le x_t(N_t)$, $y_t^*(I_t,N_t)=y_t(N_t)$, which is independent of I_t. If $I_t > x_t(N_t)$, $x_t^*(I_t,N_t)=I_t$ and, thus,

$$J_t(x_t^*(I_t,N_t),p_t,N_t)=(\bar{V}_t+\gamma(N_t)-y_t-\alpha c-b)y_t+\beta I_t+\Lambda(I_t-y_t)+\bar{\Psi}_t(I_t-y_t,\theta y_t).$$

Since $\Lambda(\cdot)$ is concave and $\Psi_t(\cdot,\cdot)$ is concave and supermodular, $J_t(x_t^*(I_t,N_t),p_t,N_t)$ is supermodular in (I_t,y_t) and its domain is a sublattice of $\mathbb{R}^2$. Hence, $y_t^*(I_t,N_t)$ is increasing in I_t for all (I_t,N_t). By Theorem 4(d), if $I_t \le x_t(N_t)$, $y_t^*(I_t,N_t)=y_t(N_t)$ is increasing in N_t. If $I_t > x_t(N_t)$, $x_t^*(I_t,N_t)=I_t$ and, thus, $J_t(x_t^*(I_t,N_t),p_t,N_t)=(\bar{V}_t+\gamma(N_t)-y_t-\alpha c-b)y_t+\beta I_t+\Lambda(I_t-y_t)+\bar{\Psi}_t(I_t-y_t,\theta y_t)$. The supermodularity of $J_t(x_t^*(I_t,N_t),p_t,N_t)$ in (y_t,N_t) follows directly from that $\gamma(\cdot)$ is increasing in N_t. Moreover, the feasible set $\{(y_t,N_t): y_t\in[\bar{V}_t+\gamma(N_t)-\bar{p},\bar{V}_t+\gamma(N_t)-\underline{p}]\}$ is clearly a sublattice of $\mathbb{R}^2$. Therefore, $y_t^*(I_t,N_t)$ is increasing in N_t for all (I_t,N_t). This proves part (d).

Part (e). If $I_t \le x_t(N_t)$, optimizing (9) yields that $\Delta_t^*(I_t,N_t)=\Delta_*$ is independent of I_t and N_t. If $I_t > x_t(N_t)$, since $I_t-\Delta_t=y_t$,

$$J_t(x_t^*(I_t,N_t),p_t,N_t)=(\bar{V}_t+\gamma(N_t)+\Delta_t-I_t-\alpha c-b)(I_t-\Delta_t)+\beta I_t+\Lambda(\Delta_t)+\bar{\Psi}_t(\Delta_t,\theta(I_t-\Delta_t)).$$

Since $\bar{\Psi}_t(\cdot,\cdot)$ is concave and supermodular, $J_t(x_t^*(I_t,N_t),p_t,N_t)$ is supermodular in (I_t,Δ_t). Moreover, the feasible set $\{(I_t,\Delta_t):\Delta_t\in[I_t-\bar{V}_t-\gamma(N_t)+\underline{p},I_t-\bar{V}_t-\gamma(N_t)+\bar{p}]\}$ is clearly a sublattice of $\mathbb{R}^2$. Hence, $\Delta_t^*(I_t,N_t)$ is increasing in I_t for all (I_t,N_t). Moreover, since $\Delta_t^*(I_t,N_t)=I_t-y_t^*(I_t,N_t)$, by part (d), $\Delta_t^*(I_t,N_t)$ is decreasing in N_t. This proves part (e). *Q.E.D.*

Appendix D: Additional Discussions on Assumption 1

In this section, we present some additional discussions on the key technical assumption of our model, Assumption 1, which assumes that $R_t(\cdot,\cdot)$ is jointly concave on its domain. Specifically, we present the necessary and sufficient conditions for $R_t(\cdot,\cdot)$ to be jointly concave in Section D.1. Then, in Section D.2, we give concrete examples of concave $R_t(\cdot,\cdot)$ functions for specific forms of the function $\gamma(\cdot)$.

D.1. Necessary and Sufficient Conditions for Assumption 1

First, we give the necessary and sufficient condition for the joint concavity of $R_t(\cdot,\cdot)$.

LEMMA 6. *Assumption 1 holds for period t, if and only if, for all $N_t \ge 0$,*

$$-2(\underline{p}-\alpha c-b)\gamma''(N_t)\ge(\gamma'(N_t))^2. \tag{59}$$

Proof: Since $\gamma(\cdot)$ is twice continuously differentiable, $R_t(\cdot,\cdot)$ is twice continuously differentiable, and jointly concave in (p_t,N_t) if and only if the Hessian of $R_t(\cdot,\cdot)$ is negative semi-definite, i.e., $\partial^2_{p_t}R_t(p_t,N_t)\le 0$, and $\partial^2_{p_t}R_t(p_t,N_t)\partial^2_{N_t}R_t(p_t,N_t)\ge(\partial_{p_t}\partial_{N_t}R_t(p_t,N_t))^2$, where $\partial^2_{p_t}R_t(p_t,N_t)=-2$, $\partial^2_{N_t}R_t(p_t,N_t)=(p_t-b-\alpha c)\gamma''(N_t)$, and $\partial_{p_t}\partial_{N_t}R_t(p_t,N_t)=\gamma'(N_t)$. Hence, $R_t(\cdot,\cdot)$ is jointly concave on $[\underline{p},\bar{p}]\times[0,+\infty)$ if and only if $-2(p_t-b-\alpha c)\gamma''(N_t)\ge(\gamma'(N_t))^2$ for all (p_t,N_t). Since $-2(p_t-b-\alpha c)\gamma''(N_t)\ge-2(\underline{p}-b-\alpha c)\gamma''(N_t)$,

$-2(p_t - b - \alpha c)\gamma''(N_t) \geq (\gamma'(N_t))^2$ for all (p_t, N_t) if and only if $-2(\underline{p} - b - \alpha c)\gamma''(N_t) \geq (\gamma'(N_t))^2$ for all $N_t \geq 0$. *Q.E.D.*

The necessary and sufficient condition for Assumption 1 characterized by inequality (59) offers little insight on what Assumption 1 means in practice. Thus, we give a simpler condition for Assumption 1.

LEMMA 7. *Let* $M := \sup\{-(\gamma'(N_t))^2/\gamma''(N_t) : N_t \geq 0\}$. $R_t(\cdot,\cdot)$ *is jointly concave if and only if* $\underline{p} \geq \alpha c + b + \frac{M}{2}$.

Proof: By Lemma 6, if Assumption 1 holds, we have $-2(\underline{p} - \alpha c - b)\gamma''(N_t) \geq (\gamma'(N_t))^2$. Since $\gamma''(N_t) \leq 0$, (59) implies that $\underline{p} - \alpha c - b \geq -\frac{1}{2}((\gamma'(N_t))^2/\gamma''(N_t))$ for any $N_t \geq 0$. Taking supreme, we have that $\underline{p} \geq \alpha c + b + \frac{M}{2}$.

If $\underline{p} \geq \alpha c + b + \frac{M}{2}$, we have $-2(\underline{p} - \alpha c - b)\gamma''(N_t) \geq -M\gamma''(N_t)$ for any N_t. Since $M = \sup\{-(\gamma'(N_t))^2/\gamma''(N_t) : N_t \geq 0\}$, $M \geq -(\gamma'(N_t))^2/\gamma''(N_t)$ for any $N_t \geq 0$. Therefore, $-M\gamma''(N_t) \geq (\gamma'(N_t))^2$ for any N_t. Putting everything together, we have $-2(\underline{p} - \alpha c - b)\gamma''(N_t) \geq -M\gamma''(N_t) \geq (\gamma'(N_t))^2$. By Lemma 6, $R_t(\cdot,\cdot)$ is jointly concave. *Q.E.D.*

D.2. Examples of Concave $R_t(\cdot,\cdot)$ Functions

We continue our discussion by giving some concrete examples of jointly concave $R_t(\cdot,\cdot)$ functions (see Assumption 1). We characterize the necessary and sufficient conditions under which $R_t(\cdot,\cdot)$ is jointly concave for some specific forms of the function $\gamma(\cdot)$. We discuss three families of $\gamma(\cdot)$: (a) exponential functions; (b) power functions; and (c) logarithm functions. We demonstrate that the necessary and sufficient conditions characterized in Lemmas 7 and 6 can be satisfied by these simple $\gamma(\cdot)$ functions under certain conditions, which are presented in model primitives and easy to verify.

First, we specify the functional form of $\gamma(\cdot)$ as $\gamma(N_t) = \gamma_0 - \gamma_0 \exp(-kN_t)$ for $N_t \geq 0$ $(\gamma_0, k > 0)$. First, we compute the first and second order derivatives of $\gamma(\cdot)$:

$$\begin{cases} \gamma'(N_t) = k\gamma_0 \exp(-kN_t), \\ \gamma''(N_t) = -k^2\gamma_0 \exp(-kN_t). \end{cases} \tag{60}$$

Note that $-\frac{(\gamma'(N_t))^2}{\gamma''(N_t)} = \gamma_0 \exp(-kN_t) \leq \gamma_0$. Hence, the necessary condition characterized in Lemma 7 for $R_t(\cdot,\cdot)$ to be jointly concave is satisfied for this family of $\gamma(\cdot)$'s. Next we characterize the necessary and sufficient condition for $R_t(\cdot,\cdot)$ to be jointly concave for an exponential $\gamma(\cdot)$ function.

LEMMA 8. *If* $\gamma(N_t) = \gamma_0 - \gamma_0 \exp(-kN_t)$ $(\gamma_0, k > 0)$, *we have* $R_t(\cdot,\cdot)$ *is jointly concave in* (p_t, N_t) *if and only if*

$$2(\underline{p} - \alpha c - b) \geq \gamma_0. \tag{61}$$

Proof: Plug (60) into (59), and we have that $R_t(\cdot,\cdot)$ is jointly concave on its domain if and only if

$$2(\underline{p} - \alpha c - b)k^2\gamma_0 \exp(-kN_t) \geq k^2\gamma_0^2 \exp(-2kN_t), \text{ for any } N_t \geq 0. \tag{62}$$

Direct algebraic manipulation yields that (62) is equivalent to that $2(\underline{p} - \alpha c - b)\exp(kN_t) \geq \gamma_0$ for any $N_t \geq 0$. Therefore, $R_t(\cdot,\cdot)$ is jointly concave if and only if (61) holds. *Q.E.D.*

Lemma 8 specifies the necessary and sufficient conditions characterized in Lemmas 7 and 6 in the case with an exponential $\gamma(\cdot)$ function. In short, $R_t(\cdot,\cdot)$ is jointly concave on its domain if and only if $\underline{p}$ is sufficiently large relative to γ_0, which is equivalent to that the price elasticity of demand is sufficiently high compared with the rating elasticity of demand.

Next, we specify the functional form of $\gamma(\cdot)$ as $\gamma(N_t) = \gamma_0 - \gamma_0(N_t+1)^{-k}$ for $N_t \geq 0$ $(\gamma_0, k > 0)$. First, we compute the first and second order derivatives of $\gamma(\cdot)$:

$$\begin{cases} \gamma'(N_t) = k\gamma_0(N_t+1)^{-k-1}, \\ \gamma''(N_t) = -k(k+1)\gamma_0(N_t+1)^{-k-2}. \end{cases} \tag{63}$$

Note that for $N_t \geq 0$, $-\frac{(\gamma'(N_t))^2}{\gamma''(N_t)} = \frac{k\gamma_0}{k+1}(N_t+1)^{-k} \leq \frac{k\gamma_0}{k+1}$. Hence, the necessary condition characterized in Lemma 7 for $R_t(\cdot,\cdot)$ to be jointly concave is satisfied. Next we characterize the necessary and sufficient condition for $R_t(\cdot,\cdot)$ to be jointly concave for a power $\gamma(\cdot)$ function.

Lemma 9. *If $\gamma(N_t) = \gamma_0 - \gamma_0(N_t+1)^{-k}$ for $N_t \geq 0$ $(\gamma_0, k > 0)$, we have $R_t(\cdot,\cdot)$ is jointly concave on its domain if and only if*

$$2(\underline{p} - \alpha c - b)(k+1) \geq \gamma_0 k. \tag{64}$$

Proof: Plug (63) into (59), and we have that $R_t(\cdot,\cdot)$ is jointly concave on its domain if and only if

$$2(\underline{p} - \alpha c - b)k(k+1)\gamma_0(N_t+1)^{-k-2} \geq k^2\gamma_0^2(N_t+1)^{-2k-2}, \text{ for any } N_t \geq 0. \tag{65}$$

Direct algebraic manipulation yields that (65) is equivalent to $2(\underline{p} - \alpha c - b)(k+1)(N_t+1)^k \geq k\gamma_0$ for all $N_t \geq 0$. Therefore, $R_t(\cdot,\cdot)$ is jointly concave on its domain if and only if (64) holds. *Q.E.D.*

Lemma 9 specifies the necessary and sufficient conditions characterized in Lemmas 7 and 6 in the case with a power $\gamma(\cdot)$ function. As in the case with exponential network externalities functions, $R_t(\cdot,\cdot)$ is jointly concave on its domain if and only if $\underline{p}$ is sufficiently large relative to γ_0, which is equivalent to that the price elasticity of demand is sufficiently high compared with the rating elasticity of demand.

Finally, we specify the functional form of $\gamma(\cdot)$ as $\gamma(N_t) = \gamma_0 \log(N_t+1)$ $(\gamma_0 > 0)$. First, we compute the first and second order derivatives of $\gamma(\cdot)$:

$$\begin{cases} \gamma'(N_t) = \frac{\gamma_0}{N_t+1}, \\ \gamma''(N_t) = -\frac{\gamma_0}{(N_t+1)^2}. \end{cases} \tag{66}$$

Note that $-\frac{(\gamma'(N_t))^2}{\gamma''(N_t)} = \gamma_0$ for all N_t. Hence, the necessary condition characterized in Lemma 7 for $R_t(\cdot,\cdot)$ to be jointly concave is satisfied for this family of $\gamma(\cdot)$'s. Next we characterize the necessary and sufficient condition for $R_t(\cdot,\cdot)$ to be jointly concave for a logarithm $\gamma(\cdot)$ function.

Lemma 10. *If $\gamma(N_t) = \gamma_0 \log(N_t+1)$ $(\gamma_0 > 0)$, we have $R_t(\cdot,\cdot)$ is jointly concave on its domain if and only if:*

$$2(\underline{p} - \alpha c - b) \geq \gamma_0. \tag{67}$$

Proof: Plug (66) into (59), and we have that $R_t(\cdot,\cdot)$ is jointly concave on its domain if and only if

$$2(\underline{p} - \alpha c - b)\frac{\gamma_0}{(N_t+1)^2} \geq \frac{\gamma_0^2}{(N_t+1)^2}. \tag{68}$$

Direct algebraic manipulation yields that (68) is equivalent to $2(\underline{p} - \alpha c - b) \geq \gamma_0$ for all $N_t \geq 0$. Therefore, $R_t(\cdot,\cdot)$ is jointly concave on its domain if and only if (67) holds. *Q.E.D.*

Lemma 10 specifies the necessary and sufficient conditions characterized in Lemmas 7 and 6 in the case of power $\gamma(\cdot)$ functions. As in the cases with exponential and power $\gamma(\cdot)$ functions, $R_t(\cdot,\cdot)$ is jointly concave on its domain if and only if $\underline{p}$ is sufficiently large relative to γ_0, which is equivalent to that the price elasticity of demand is sufficiently high compared with the rating elasticity of demand.

Lemmas 8-10 confirm our previous insight delivered by Lemma 7 that when the price elasticity of demand (i.e., $|\frac{\mathrm{d}\mathbb{E}[D_t(p_t,N_t)]/\mathbb{E}[D_t(p_t,N_t)]}{\mathrm{d}p_t/p_t}|$) is sufficiently high relative to the *rating elasticity of demand* (i.e., $|\frac{\mathrm{d}\mathbb{E}[D_t(p_t,N_t)]/\mathbb{E}[D_t(p_t,N_t)]}{\mathrm{d}N_t/N_t}|$), $R_t(\cdot,\cdot)$ is jointly concave in (p_t,N_t) on its domain. Therefore, Assumption 1 can be satisfied for a wide variety of $\gamma(\cdot)$ functions. To conclude this section, we remark that the above method can be easily adapted to characterize the conditions under which $R_t(\cdot,\cdot)$ is jointly concave with other families of $\gamma(\cdot)$ functions.

Appendix E: Parameters for the Numerical Illustrations in Figures 1-2

This section offers the parameter specifications for the numerical illustrations in Figures 1-2, which plot the behaviors of the optimal policy throughout the planning horizon for different values of the aggregate net rating. The numerical example belongs to the family of numerical experiments examined in Section 5.3. The parameters are given in Table 2.

Table 2 Parameter Specifications for Figures 1-2

$\bar{V}_t = 30$:	Stationary market throughout the planning horizon
$T = 10$:	planning horizon length
$\gamma(N_t) = 0.5N_t$:	impact of aggregate net rating on demand
$D_t(p_t,N_t) = 30 + 0.5N_t - p_t + \xi_t$:	demand function
$\xi_t \sim N(0,2)$:	demand perturbation follows a truncated normal distribution to ensure $D_t(p_t,N_t) \geq 0$
$\alpha = 0.99$:	discount factor
$c = 8$:	inventory purchasing cost
$b = 10$:	backlogging cost
$h = 1$:	holding cost
$[\underline{p},\bar{p}] = [0,25]$:	price range
$\eta = 0.5$:	discount factor for reviews
$\theta = 0.5$:	net rating contribution ratio of demand
$\sigma = 0$:	impact of inventory availability on net rating

Appendix F: Micro-Foundation for the Paid-Review Strategy

In this section, we provide the micro-foundation for the model with the paid-review strategy. Assume that the firm sends conditional cash rewards or coupons to Z customers who recently purchased the product without leaving any review. If a customer opts to leave a review, she will receive an equivalent monetary reward of w. A customer is characterized by his willingness-to-review $\omega \geq 0$, which follows a continuous distribution with CDF $\Omega(\cdot)$ satisfying the log-concave property (i.e., $\log(\Omega(\cdot))$ is concave, see Bagnoli and Bergstrom, 2006). Hence, a customer would choose to leave a review if and only if $\omega \leq w$. We further assume that, for different customers, ω are *i.i.d.*. To ensure the effectiveness of paid-review strategy, the firm only sends cash rewards/coupons to customers who receive the product immediately. Conditioned on that a customer chooses to leave a review, i.e., $\omega \leq w$, a customer would leave a positive review with probability ς^+ and a negative review with probability ς^-, where $\varsigma := \varsigma^+ - \varsigma^- > 0$.

We are now ready to compute the function $c_n(\cdot)$. Given the monetary reward w, the total number of customers who opt to leave a review is $n_t = \varsigma Z\mathbb{P}(\omega \leq w) = \varsigma Z\Omega(w)$. Hence, $w = \Omega^{-1}\left(\frac{n_t}{\varsigma Z}\right)$ and the total cost of the paid-review strategy increase n_t aggregate net rating is given by:

$$c_n(n_t) = Z\mathbb{P}(\omega \leq w)w = \frac{n_t}{\varsigma}\Omega^{-1}\left(\frac{n_t}{\varsigma Z}\right) \tag{69}$$

Since the distribution of ω, $\Omega(\cdot)$, satisfies the log-concave property, we can show that $c_n(\cdot)$ is continuously differentiable and convexly increasing in n_t.

LEMMA 11. *Assume that $\Omega(\cdot)$ satisfies the log-concave property. Then $c_n(n_t)$, defined by* (69)*, is convexly increasing in n_t.*

Proof. The continuous differentiability of $c_n(\cdot)$ follows immediately from that ω follows a continuous distribution. Let $h(x) := \log\Omega(x)$. Since $h(x)$ is concave, we have $h'(x) = \frac{\Omega'(x)}{\Omega(x)}$ is decreasing in x (equivalently, $\frac{\Omega(x)}{\Omega'(x)}$ is increasing in x). Therefore,

$$c_n'(n_t) = \frac{1}{\varsigma}\Omega^{-1}\left(\frac{n_t}{\varsigma Z}\right) + \frac{1}{\varsigma}\cdot\frac{\Omega\left(\Omega^{-1}\left(\frac{n_t}{\varsigma Z}\right)\right)}{\Omega'\left(\frac{n_t}{\varsigma Z}\right)} > 0,$$

which is increasing in n_t, given that both $\Omega^{-1}(x)$ and $\frac{\Omega(x)}{\Omega'(x)}$ are increasing in x. This proves that $c_n(\cdot)$ is convexly increasing in n_t. *Q.E.D.*

References

[1] Bagnoli, M., T. Bergstrom. 2006. Log-concave probability and its applications. *Economic Theory.* **26**(2) 445-469.

[2] Boyd, S., L. Vandenberghe. 2004. *Convex Optimization.* Cambridge University Press, New York.

[3] Durrett, R.. 2010. *Probability, Theory and Examples, 4th Edition.* Cambridge University Press, New York.

[4] Stokey, N., R. Lucas, and E. Prescott. 1989. *Recursive Methods in Economic Dynamics.* Harvard University Press, Cambridge.

[5] Topkis, D.. 1998. *Supermodularity and Complementarity.* Princeton University Press, Princeton.